

NAVSEA ACQUISITION TEST & EVALUATION HANDBOOK



**A Definitive Source for Effective T&E Management
Of Navy Acquisition Programs**

**Published by the Naval Sea Systems Command
Test & Evaluation Office
December 2000**

Forward

The Naval Sea Systems Command (NAVSEA) Acquisition Test and Evaluation (T&E) Handbook is a guide prepared specifically for the T&E managers T&E in the ship and system acquisition programs of NAVSEA and its affiliated Program Executive Offices (PEOs). It was developed by the NAVSEA T&E Office, based on their 25 years' experience with such programs.

This update reflects the restructuring of the Acquisition DoD 5000.2 policies and procedures of October 2000. More importantly, this handbook describes the unique T&E strategies that evolved at the end of the 1990s, and which have become today's accepted practices. This Handbook will continue to evolve as new approaches to T&E are developed and implemented by managers of major and non-major acquisition programs.

We believe the discussions and guidance presented herein, based on many years of lessons learned, will assist in planning and executing T&E in support of the development of today's complex naval platforms, weapons, and combat systems.

Matthew T. Reynolds
Director, NAVSEA T&E Office

The NAVSEA T&E Office RDT&E TEAM

George Axiotis
Andy Kristovitch
Stu Rednor
Paul Hagan

Questions or comments concerning this Handbook can be forwarded to:

Commander, Naval Sea Systems Command
Test and Evaluation Office (91T)
2531 Jefferson Davis Highway
Arlington VA 22242-5160

Or reach us at: testevaluation@navsea.navy.mil

Visit the NAVSEA T&E Homepage at:
www.navsea.navy.mil/navysea-te

Table of Contents

Description	Page
Chapter One	
The Acquisition Process and T&E	1-1
1.0 Introduction	1-1
1.1 Sources of an Acquisition Program	1-1
1.2 Program Phases	1-2
1.3 Principal Program Decisions	1-3
1.4 Acquisition Categories	1-5
1.5 Test and Evaluation (T&E)	1-5
1.6 Developmental T&E (DT&E)	1-5
1.7 Operational T&E (OT&E)	1-7
1.8 OPTEVFOR	1-7
1.9 T&E Oversight	1-7
1.10 Performance Based Acquisition	1-7
1.11 T&E and Evolutionary Acquisition	1-8
1.12 Non-Developmental Items	1-8
1.13 TECHEVAL and OPEVAL	1-9
1.14 Unique T&E Requirements	1-10
1.14.1 Explosives/Energetics	1-10
1.14.2 Weapon Systems	1-10
1.14.3 Sensor Systems	1-12
1.14.4 Combat Control Systems	1-12
1.14.5 C4ISR Systems	1-13
1.14.6 Software Upgrade Programs	1-14
1.14.7 Hull, Mechanical & Electrical (HM&E) Systems	1-14
1.14.8 Ship Acquisition Programs	1-14
1.15 Live Fire T&E (LFT&E)	1-15
Chapter Two	
T&E Program Planning	2-1
2.0 Introduction	2-1
2.1 Effective T&E Planning	2-1
2.2 Managing T&E	2-2
2.3 T&E Program Team	2-2
2.4 OPTEVFOR Involvement	2-4
2.5 Requirements Document	2-3
2.6 Performance Requirements	2-3
2.7 Number of Test Articles	2-5
2.8 Efficient T&E Planning	2-5
2.8.1 Keeping Test Phases Manageable	2-6
2.8.2 "Piggyback" T&E Events	2-6
2.8.3 Harmonizing DT and OT	2-6
2.9 Modeling and Simulation (M&S)	2-7
2.10 Interoperability and T&E	2-7
2.11 Test Site Selection and Constraints	2-9
2.12 Land-Based Testing	2-9
2.13 LBTS Role in Interoperability	2-11
2.14 Specialized T&E Resources	2-11
2.15 Threat Target Validation	2-12

2.16	Designing For Testability	2-12
2.17	Data Collection and Analysis	2-12
2.18	Environmental Qualification Testing (EQT)	2-13
2.19	Reliability, Maintainability and Availability (RM&A)	2-13
2.20	Computer Software Testing	2-14
2.21	NEPA and Environmental Compliance	2-15
2.22	Operations and Data Security (OPSEC)	2-15
2.23	T&E Funding	2-16
2.24	Fleet Services and Support	2-16
2.25	Supporting Milestones and Decision Points	2-17

Chapter Three

Documenting T&E

3-1

3.0	Introduction	3-1
3.1	Test and Evaluation Master Plan (TEMP)	3-1
3.1.1	Purpose	3-1
3.1.2	Responsibilities for Preparation	3-1
3.1.3	The TEMP IPT	3-1
3.1.4	Timely Submittal and Updates	3-2
3.2	Format and Watch Areas	3-2
3.2.1	Cover Page	3-2
3.2.2	Part I	3-2
3.2.3	Part II	3-4
3.2.4	Part III	3-5
3.2.5	Part IV	3-5
3.2.6	Part V	3-5
3.3	Abbreviated Acquisition Programs (AAP)	3-5
3.4	Review, Processing and Approval	3-5
3.5	Other T&E Plans	3-6
3.6	Test Reporting	3-6
3.7	Updating Program Documents	3-7

Chapter Four

Early T&E

4-1

4.0	Introduction	4-1
4.1	Defining the Concept	4-1
4.2	Early T&E	4-1
4.3	OPTEVFOR's Involvement	4-2
4.4	Early Operational Assessments	4-2
4.5	Early Development Examples	4-2
4.6	Test and Evaluation Identification Number (TEIN)	4-3
4.7	The TEMP's Role	4-4
4.8	Assessing Requirements	4-5
4.9	Solidifying Thresholds	4-5
4.10	Threshold Pitfalls	4-6
4.11	Exiting the Initial Phase	4-7

Chapter Five

T&E During System Development, TECHEVAL and OPEVAL

5-1

5.0	Introduction	5-1
5.1	The Engineering Development Model (EDM)	5-1
5.2	Testing During System Integration	5-2

5.3	Role of the DT Report	5-2
5.5	The Interim Progress Review	5-2
5.6	System Development T&E	5-3
5.7	System Development Program Examples	5-5
5.8	Milestone C	5-7
5.9	Shipboard Installation	5-7
5.10	TECHEVAL	5-8
5.10.1	Scope of TECHEVAL	5-8
5.10.2	TECHEVAL Plan	5-9
5.10.3	TECHEVAL Report	5-9
5.11	OT Readiness	5-9
5.11.1	Readiness Criteria	5-10
5.11.2	Getting Ready for OT	5-11
5.11.3	OT Readiness Review	5-12
5.11.4	OTRR Brief	5-12
5.11.5	Certification of Readiness	5-13
5.11.6	De-Certification/Re-Certification	5-13
5.12	OPEVAL	5-13
5.12.1	OPEVAL Test Plan	5-13
5.12.2	Resources	5-14
5.12.3	Conduct of OPEVAL	5-14
5.12.4	OPEVAL Report	5-14
5.13	Full-Rate Production (FRP) Review	5-15

Chapter Six

T&E During Rate Production and Deployment

6-1

6.0	Introduction	6-1
6.1	Developmental T&E (DT III)	6-1
6.2	Follow-on Operational T&E (FOT&E)	6-1
6.2.1	Certification of Readiness for FOT&E	6-1
6.3	Technology Insertion and OT	6-1
6.4	Production Acceptance T&E (PAT&E)	6-2
6.4.1	PAT&E for Control in Production	6-2
6.4.1.1	Manufacturing Screening Tests	6-2
6.4.1.2	Burn-in Tests	6-2
6.4.1.3	Random Sampling Tests	6-2
6.4.1.4	In-process Tests	6-2
6.4.1.5	Environmental Tests	6-2
6.4.1.6	Reliability and Maintainability Demonstration Tests	6-2
6.4.1.7	Proof Tests	6-3
6.4.1.8	Electromagnetic Interference Tests	6-3
6.5	PAT&E for Control in Installation and Operation	6-3
6.6	In-service T&E	6-3
6.7	Assessment Continuation	6-3

Figures

1-1	Acquisition Process Continuum	1-4
1-2	Notional Weapon System Development	1-4
1-3	T&E Within Evolutionary Acquisition	1-9
2-1	Threat Target Validation Process	2-13
3-1	Notional TEMP Part II Integrated Schedule	3-4
3-2	TEMP Development and Approval Process	3-6
4-1	Format for TEIN Request	4-4
5-1	Test Planning Matrix (Example)	5-4
5-2	OT Readiness Timeline	5-12

Appendix A – OT Readiness Checklist**A-1****Appendix B – T&E Manager’s Tool Kit****B-1****Appendix C – Index****C-1**

CHAPTER ONE

THE ACQUISITION PROCESS AND T&E

1.0 Introduction

Test and Evaluation (T&E) plays a critical role in the engineering of ships, weapons, and combat systems. It is the primary vehicle by which performance risks are identified and quantified. As such, T&E is an important management discipline within the defense acquisition process. T&E results are a key determinant in program approvals, and acquisition programs must be structured to ensure that meaningful T&E results are available at the key decision points in each program. The directives issued by the Office of the Secretary of Defense (OSD) and the Navy place special emphasis on this role of the T&E program. However, the objective of T&E is not just to confirm, but also to learn. T&E is an integral part of the systems engineering that is key to the successful development of ships and systems that meet the needs of the Fleet. This handbook describes processes, procedures and best practices that have proven to be effective in managing T&E programs in NAVSEA and its affiliated PEOs.

The objective of T&E is not just to confirm, but to learn.

The NAVSEA Test and Evaluation Office provides this Handbook as a primary source of information for managing complex ship, weapon, and combat system T&E efforts. It is meant to bridge the gap between the policy provided in the directives and knowledge and publications that exist in the Navy engineering organizations and the industry that supports us.

The objective of confirming adequate performance is affirmed by policy that has its roots in Public Law. Title 10 United States Code (US Code) Section 2399 requires that an independent phase of Operational T&E (OT&E) be conducted before full rate production. Likewise, Section 2366 of Title 10 states that Live-Fire T&E is to be conducted (in programs to which it applies) before full rate production. With such "legal" emphasis on the pass/fail dimension of T&E, there are pressures to focus so much attention on achieving the quantified performance thresholds that the other thrusts of T&E would suffer. Accelerating or reducing testing primarily for the sake of meeting a previously defined acquisition program checkpoint will usually only cost

more in time and funding the long run, and may even take a toll on the customers' confidence in the acquisition process. This chapter describes this process in terms that set the stage for describing the tenets of the management of T&E programs.

1.1 Sources of an Acquisition Program

The need for an acquisition program arises from a variety of sources, such as recognition of an performance shortfall in an existing system, a requirement to establish new capabilities to meet a new threat, a decision to capitalize on a technologically feasible opportunity to lower life-cycle cost or total ownership cost, or the necessity of a change to strategic or tactical doctrine. In line with the OSD emphasis on limiting the technology risks bought forward as part of acquisition programs, a formal acquisition program is said to begin upon successful approval at Milestone B for the new DoD Process and Milestone I under the legacy process. The efforts prior to these points are considered concept and technology development. The start of an acquisition program is dependent on three things: technology (including software) maturity, validated requirements, and funding.

Acquisition programs include not only totally new systems, but also significant upgrades to existing in-service systems. Systems that are procured "off-the-shelf" (i.e., items commercially available and bought "as is," without significant modifications required by the government) are not usually managed as acquisition programs, unless they are designated as "Abbreviated ACAT". For example, a commercially available network switch is found to be a suitable functional replacement for existing shipboard MIL-SPEC equipment. In this case, additional development work and T&E may be required to address the potential risks of using this device in a shipboard environment. SECNAVINST 5000.2B Section 1.4.3 (<http://www.deskbook.osd.mil>) outlines the decision process by which modifications to existing systems are assessed for assignment as an acquisition program. Assistant Secretary of the Navy for Research, Development and Acquisition. (ASN(RDA)), in consultation with the OPNAV program sponsor, the OPNAV T&E Division (CNO N912), the PM, and others, assesses the scope and impact of proposed program. In cases where

a formal acquisition program is warranted, an ACAT level (para 1.4) is assigned by ASN (RDA).

Systems developed in acquisition programs evolve from multiple sources. Totally new systems are, at times, derived from basic technology; however, a program normally consists of a combination of developing new technology and using applications from other sources, such as: commercial off-the-shelf systems, Foreign Weapons Evaluation (FWE) programs, in-service Navy systems, other services' existing or developmental systems, and Rapid Prototyping programs. The sources in a particular program determine the risks and uncertainties, which, in turn, are drivers for the scope and depth of the T&E program. SECNAVINST 5000.2 contains a logic diagram for determining when systems can be designated as ACAT programs.

1.2 Program Phases

There are three basic evolutionary steps to complex system engineering development from which the Defense Acquisition process is based. They are; Subsystem/technology maturation, System Development and Production readiness. OSD has defined these three aspects into phases of an acquisition program that reflect the maturation of a system from concept to fielding. The process since the early 1980's, known throughout here as the "Legacy" process, identifies four phases of development; Phase 0 - Concept Exploration (CE), Phase II - Program Definition and Risk Reduction (PDRR), Phase III - Engineering & Manufacturing Development (EMD), and Phase IV - Production, Fielding/Deployment. A new acquisition milestone labeling was introduced in FY01 that is applicable to all Acquisition programs initiated in, or after that fiscal year. The process, known here within as the "New" process. This process outlines only three phases of an acquisition program; Phase I - Concept & Technology Development (CAD), Phase II - System Development & Demonstration (SDD), and Phase III - Production & Deployment (P&D). The difference in these two processes centers around when a formal program begins and when production is initiated. Such difference in phase delineation does not substantially add or detract the amount or type of T&E that is conducted. Regardless of the process used, each of these phases is separated by discrete decision points known as Milestones. T&E is one of the major aids supporting these decision points. Figure 1-1 on the following page shows how these phases are tied to a notional program schedule.

- **Concept and Technology Development.** This phase of development has two primary objectives: to solidify the concept and identify the mature technologies that can be used; and to validate the system concept and demonstrate the feasibility of executing engineering, manufacturing, and development of the proposed system. Known as the combination of Phases 0 and I for the *Legacy* process and Phase I of the *New* Process, it

identifies alternative system concepts to satisfy the need and selects those to be carried forward into System Development. Alternatives are evaluated through studies and demonstrations such as the Analysis of Alternatives (AOA), Advanced Technology Demonstrations (ATD), Concept Demonstrations, and Fleet Battleforce Experiments. The initial operational concepts and requirements are formulated. The desired concept is then selected. T&E is generally limited to subsystem breadboard models, simulation, and modeling results. The second aspect of this period involves: (1) refinement of competing alternative system concepts; (2) assessment of program risk and uncertainties; (3) refinement of cost estimates; and (4) the construction and testing of an Advanced Development Model (ADM) and refinement of technical and operational performance requirements through a series of Advanced Concept Technology Demonstrations (ACTD). If the risk is small, this phase, or phases in the case of the *Legacy* process, may be omitted and the program can proceed directly to the next step. T&E near the end of this technology and subsystem maturation is limited to prototypes of varying degrees of maturity demonstrated at-sea or at a land-based facility, to a limited degree. The results are compared to additional modeling and simulation previously conducted.

- **System Development.** This step essentially defines most acquisition programs beginning at Milestone II (*Legacy*) and Milestone B (*New*). The objective is to design and develop the system, along with its logistic support. A limited number of Engineering Development Models (EDMs) for development and T&E is procured, and cost estimates and performance and schedule thresholds are reviewed for consistency with the risks involved. Units may also be acquired under a Low Rate Initial Production (LRIP) contract to supply additional articles for T&E. The vast majority of T&E is conducted here and is much more formalized involving many resources, and is treated as a management discipline.

- **Production and Deployment.** The objective of the production and deployment is to transition a system from development into limited-rate initial production, followed by full-rate production. A Low-Rate Initial Production decision may be made during the early portion of this process to provide units in support of T&E. Full Rate Production is authorized through a later Decision Review (Milestone III under the *Legacy* process) conducted by the Milestone Decision Authority (MDA) or an official to whom the MDA

has delegated that authority. Planning for deployment, including the assignment and training of personnel and establishment of logistics support, is required. During this period, fully integrated systems undergo final developmental T&E and operational T&E, and are delivered to the Fleet.

Each T&E program must be tailored to the risks and uncertainties inherent in the acquisition program it supports. An overall T&E strategy must be defined, and must include the scope and tasks for each phase, the particular tests to be accomplished, the funding profile, and the procurement approach.

1.3 Principal Program Decisions

Prior to the start of each phase in an acquisition program, there is a major review to reassess: the need for the system, the estimated ability of that system to counter known and perceived threats, whether technological advances are being achieved as planned, program schedules, and the prospects for funding the continuation of the program. These reviews are known as "Milestone Reviews." After each review, a decision is made to move ahead as planned to modify the plan, or to discontinue the program.

Figure 1-1 depicts the Acquisition Process. Figure 1-2 represents a notional major ACAT weapon program requiring Live-Fire testing (LFT&E), and illustrates the basic process and major decision points established for the acquisition process.

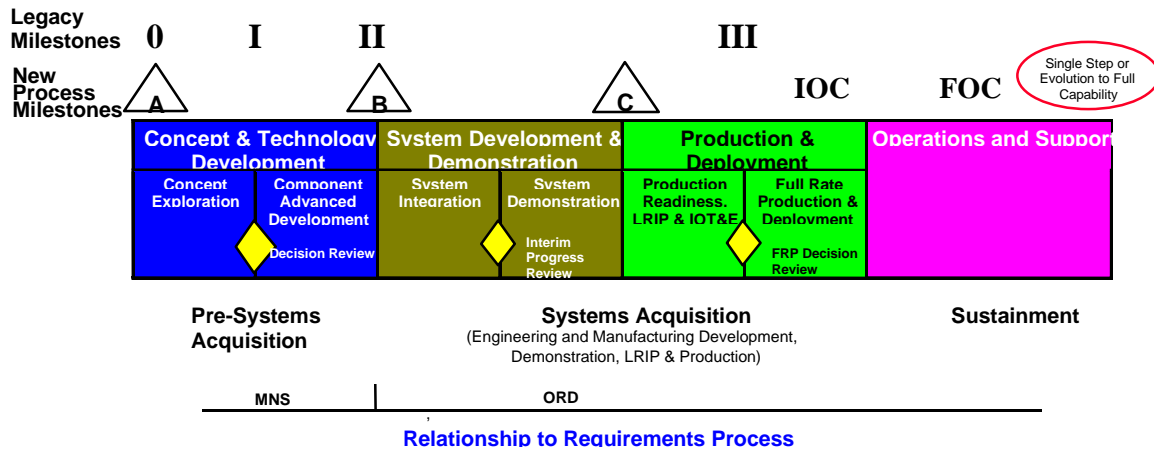


Figure 1-1 Acquisition Process Continuum

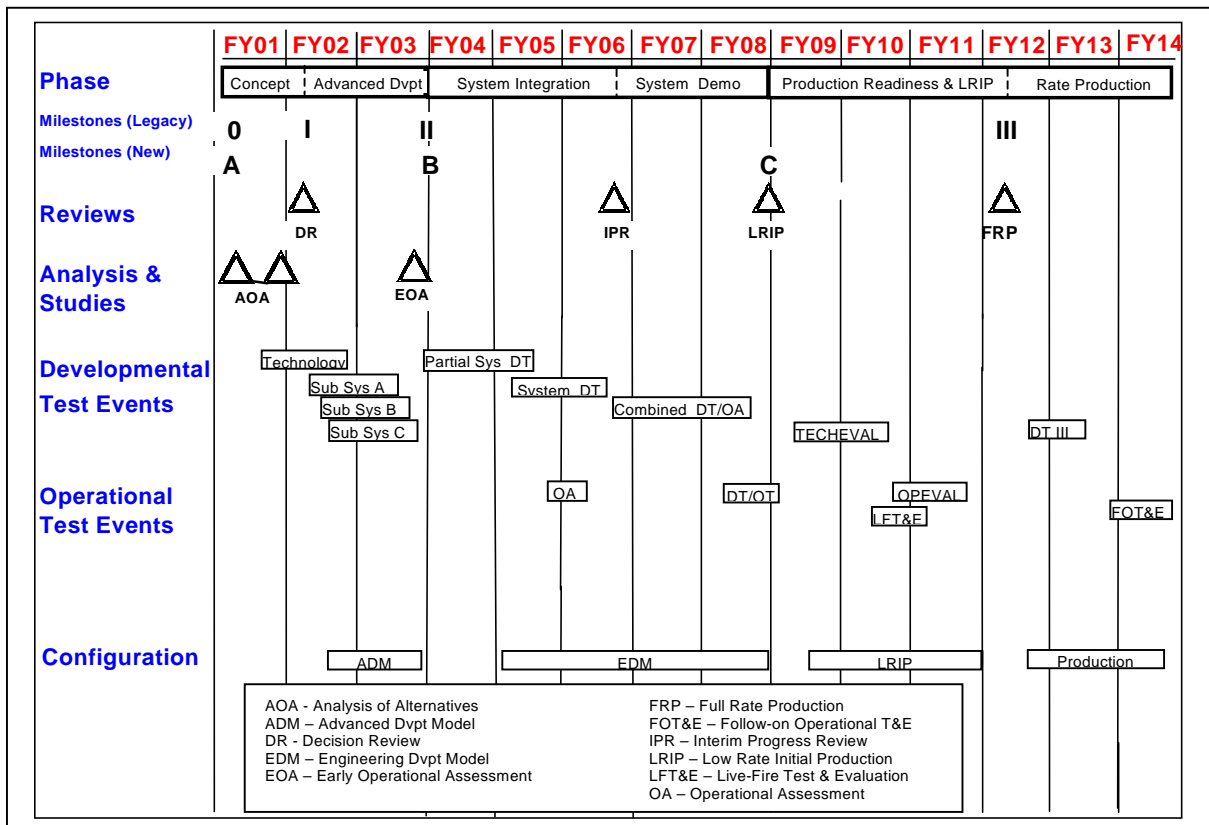


Figure 1-2 Notional Weapon System Development

1.4 Acquisition Categories

In order to focus higher management attention on higher priority programs, DoD has established four Acquisition Categories (ACATs). The Navy has expanded this to six categories, each with a different echelon assigned as the final authority at the milestone reviews. The Assistant Secretary of the Navy for Research Development and Acquisition (ASN (RDA)) publishes the ACAT assignment of Navy Acquisition programs in a semiannual Acquisition Program Listing. The following is a breakdown of ACATs and the typical percentage of programs in each. Approximately 125 ACAT programs are managed at any one time by NAVSEA and its affiliated Program Executive Officers (PEO).

ACAT	Production Decision Authority	Percentage
ID	UNSECDEF (Acquisition)	5
IC	ASN (RDA)	10
II	ASN (RDA)	10
III	ASN (RDA) or PEO	35
IVT	SYSCOM or PEO	22
IVM	SYSCOM or PEO	18

Generally, ACATs are assigned according to the overall cost of the program (both R&D and production). However, for ACAT III and IV, cost is not a determining factor. The primary distinction between ACAT III and IV programs is that those systems that significantly and directly affect the Navy's combat capability (i.e., have relatively direct interaction with the enemy) are ACAT III, and those that do not are ACAT IV. For the Navy, ACAT IV programs are further divided into ACAT IVTs [including Operational T&E conducted by the Operational Test and Evaluation Force (OPTEVFOR)] (discussed in the following paragraphs) and ACAT IVMs [not including T&E conducted by OPTEVFOR]. If OPNAV believes that an ACAT IV program should have Operational T&E before a production decision is made, it will be designated IVT.

***T&E results are to be a pacing item
in the progress of the program.***

1.5 Test and Evaluation (T&E)

DoD Directive 5000.2 (www.deskbook.osd.mil), a fundamental document that describes acquisition policies, stipulates that T&E is to have a major input to acquisition decision making, and is to be a pacing item in the progress of the program. This document also defines the key terminology and establishes DoD policy for the conduct of T&E. It requires that T&E strategy, results to date, and future planning be part of the milestone reviews. The instruction also has separate Appendices on the TEMP and LFT&E.

Two types of T&E are conducted during the acquisition process: Developmental T&E (DT&E) and Operational T&E (OT&E).

Developmental Test and Evaluation (DT&E) is T&E conducted to assist the engineering and design process in verifying progress towards achievement of technical performance specifications and objectives. DT is conducted by the developing agencies (DA), which include contractor(s), sub-contractor(s), Navy labs, and Navy field and at-sea activities for the program manager. DT constitutes the vast majority of the T&E conducted during system acquisition.

Operational Test and Evaluation (OT&E) is T&E conducted by Commander, Operational Test and Evaluation Force (COMOPTEVFOR) to estimate a system's operational effectiveness and operational suitability. (See paragraph 2.5 for a definition of these two terms.) Each Service has an agency, independent of the user and the producer commands, responsible for planning, conducting, and reporting OT&E. The Navy's OT&E agent is OPTEVFOR. OT&E is subdivided into Initial Operational T&E (IOT&E), which is all OT&E, up to and including the Operational Evaluation (OPEVAL) (i.e., generally the end of R&D) and Follow-on Operational T&E (FOT&E), which is all OT&E after the final phase of OPEVAL (after the full-rate production decision).

A given system acquisition program usually contains an optimum mix of and schedule for the two types of T&E, tailored to reduce the performance risks and provide concrete evidence that the program is achieving the anticipated technical progress. The objectives and types of T&E will vary, depending on the phase of the program, its acquisition strategy, and the technical risks.

1.6 Developmental T&E (DT&E)

The systematically planned actions that transform an operational need into a description of system performance parameters and a preferred system configuration is known as systems engineering. The program manager's job is to direct the Navy and industry teams' systems engineering efforts in the planning and control of technical program tasks; the integration of the engineering specialties; and the integration of design, test, logistics, and production engineering to meet cost, schedule, and technical performance requirements. Developmental T&E is a subset and an integral part of systems

engineering in an acquisition program. The phasing of DT&E indicates how DT&E objectives change during the course of a program.

DT-I is DT&E conducted in support of concept and technology maturation and subsystem development to demonstrate that design risks have been identified and minimized. It is normally conducted at the subsystem/component level, up to and including employment of advanced development models for final evaluation. Design, operation, and implementation risks are mitigated during this phase through a series of DT-I Technology and Concept Demonstrations.

DT-II is DT&E conducted during system development to support the production decision. DT-II demonstrates that the design meets the requirements for performance, reliability, maintainability, operational availability, logistic support, survivability, vulnerability, compatibility, interoperability, human factors, transportability, safety, training, and electromagnetic environmental effects. The final phase of DT II is the TECHEVAL, conducted to verify the achievement of technical performance goals and assess readiness for OPEVAL. Section 1.10 and Chapter 5 contain a complete description of TECHEVAL.

DT-III is DT&E conducted after the production and deployment decision to verify that product improvements, or correction of design deficiencies discovered during R&D, are effective.

1.7 Operational T&E (OT&E)

Operational Test and Evaluation (OT&E is approached from a different perspective than DT&E. DT&E is characterized as a subset of engineering, where the test requirements are traceable to design and performance requirements, and production acceptance tests are traceable to production specifications.

Navy OT&E is the sole domain of OPTEVFOR. If OPTEVFOR is not involved in a particular test event, it cannot, by definition, be OT&E.

Operational T&E has a different objective. The need for operational testing has its basis not in the engineering process of building, testing, fixing, and re-testing; rather, OT&E resulted from the need to test a system in its operating environment. DoD systems are operated and maintained in environments that can have seriously limiting or degrading effects on their performance. Such effects are not always readily uncovered during DT. For that reason, the Defense Department, including the Navy, requires the conduct of rigorous operational testing during the R&D program, as well as the conduct of at least some user-oriented, at-sea

technical testing in preparation for the operational testing events. Enough risk exists in placing a new system in its operational environment that the final decision on whether or not the system will be mass-produced and deployed is delayed until operational T&E can be conducted. That is one the reasons why an independent activity, Commander, Operational Test and Evaluation Force (COMOPTEVFOR), separated organizationally from the Developing Agencies, is assigned to conduct this operational T&E. Special note should be made that Navy OT&E is the sole domain of OPTEVFOR. If OPTEVFOR is not involved in a particular test event, it cannot, by definition, be OT&E. Initial Operational Test and Evaluation (IOT&E) is all OT&E conducted before the full-rate production decision. Follow-on Operational Test and Evaluation (FOT&E) is all OT&E after the full-rate production decision. The phases of OT&E are classified according to the Acquisition phase in which they are conducted:

OT-I is IOT&E conducted to provide an early estimate of projected operational effectiveness and operational suitability of the system, initiate tactics development, estimate program progress, and identify operational issues for OT-II. OT-I is required in relatively few programs.

OT-II is IOT&E conducted for two reasons. First, it is conducted to identify weaknesses in the system, and make operational effectiveness predictions of systems as they mature through System Development. Secondly, OT is conducted to demonstrate the achievement of program requirements for operational effectiveness and operational suitability, and continuing tactics development in support of the Full Rate Production decision. The final phase of OT II is the Operational Evaluation (OPEVAL), conducted to verify achievement of technical performance goals and assess the ability of the system to function in an operational environment using fleet sailors as required by Law.

OT-III is FOT&E generally conducted with the same pre-production prototype or pilot production systems used in OT-II. It may also be conducted on the first production units to assess configuration differences between the OPEVAL (the final phase of OT-II) and production configurations. Specific OT-III objectives include testing of fixes to be incorporated in production systems, completion of any deferred or incomplete IOT&E, and continuing tactics development. For selected ship acquisition programs, OT-III is conducted with the lead ship

during the period between delivery and expiration of SCN funding authority.

1.8 OPTEVFOR

The Operational Test and Evaluation Force (OPTEVFOR) (<http://www.cotf.navy.mil>) is a Fleet-oriented command, staffed largely by officers who have had recent operational duty. They are chartered to use their operational expertise to conduct a series of tests to verify a system's operational effectiveness and suitability. The OPTEVFOR Force is provided the resources necessary to operationally test these systems via the program office. OPTEVFOR does not limit the scope of their evaluation to just the unit under test. Their charter is to examine the system as a whole, as well as its ability to operate with other systems. Other resources needed to perform operational testing, such as Fleet ships and aircraft, range time, targets, and Fleet sailors, are provided for or authorized by the CNO.

OPTEVFOR is an operational Fleet command with policy direction, technical and procedural guidance, and financial support coming from the Chief of Naval Operations (CNO) to whom he reports directly. For the operation of Fleet units, COMOPTEVFOR reports to CINCPACFLT, CINCLANTFLT, and when necessary, CINCUSNAVEUR. COMOPTEVFOR's headquarters are in Norfolk, Virginia. OPTEVFOR maintains two aircraft squadrons and one detachment: Air Test & Evaluation Squadron One (VX-1), Naval Air Station, Patuxent River, Maryland; Air Test & Evaluation Squadron Nine (VX-9), Naval Weapons Center, China Lake, California; and Air Test & Evaluation Squadron Detachment Nine (VX-9Det), Naval Air Station, Point Mugu, California.

OPTEVFOR obtains the services of "Trusted Agents" for those programs requiring specialized analysis, or when the nature or scope of testing necessitates on-site test support. The role of the Trusted Agent (TA) is outlined in their own Operational Test Director's (OTD) Guide (<http://www.cotf.navy.mil/otd/cover2.htm>), and describes the process by which the TA learns the system early on and then helps OPTEVFOR assess it during OT and OA events. The PM is required to question the selection of the TA, and may propose a different arrangement to OPTEVFOR if deemed necessary.

OPTEVFOR provides a series of reports and assessments to support the acquisition process. They are tailored in scope and complexity to support the acquisition process and the needs of the Milestone Decision Authority (MDA) and program manager (PM). These are: Early Operational Assessments (EOA) – Concept and Technology analysis of risks, Operational Assessments (OA) – Reports of early testing within System Development on immature systems, Operational Test (OT) reports – Formal assessment supporting the Full Rate Production decision and Development Test

Assist (DT Assist) summaries – low-level observations of DT supporting the program manager, Observation of Operational Capability (OOC) – low-level observation of technology insertion, or new capability demonstrations in the Production Phase. The PM works closely with OPTEVFOR to determine the most appropriate reports needed to support the acquisition strategy.

1.9 T&E Oversight.

About one fourth of NAVSEA/PEO acquisition programs are under OSD Oversight. They are primarily ACAT I and II programs, but even a few ACAT III and IV programs have been designated by OSD for oversight. Oversight essentially means that those programs must obtain OSD approval of their Test and Evaluation Master Plans (TEMP) and Operational T&E plans. The PM can expect to have extensive participation of OSD staff in T&E planning working groups. For those programs designated Live-Fire T&E (LFT&E) programs, the LFT&E Management plan must be coordinated and approved through the Director of Operational T&E (DOT&E) on the OSD staff.

1.10 Performance Based Acquisition

In most R&D programs, it is rare that program managers procure a data package (specifications and drawings) that would ensure that what was tested during development was in fact what was procured in production. For Commercial-Off-The-Shelf (COTS) and Non-Developmental Item (NDI) procurements structured around performance-based requirements and specifications, the trend has been to forgo procurement of a data package and transfer all configuration control to the vendor. From a T&E perspective, the units manufactured during Production and Deployment will most likely differ from those tested in TECHEVAL and OPEVAL. New configurations, whether instituted by the original developer or a new contractor entering the program during production, can invalidate much of the T&E results achieved during development. For example, a change in components can greatly impact spare parts support for the system and thereby seriously degrade the operational availability the system achieves when deployed. Program managers should be aware of such pitfalls. With emphasis on technology infusion using COTS/NDI, there is a strong possibility that the units produced after the initial lot may differ. Even more critical is that the manufacturer may not have proper documentation on the types of testing that may have been conducted on these sub-systems. Thus, there will always be need for additional T&E, either as a formal production T&E program,

an in-service assessment by the fleet or OPTEVFOR to assess compliance with the original requirements and to identify interface issues with the new configuration.

Evolutionary acquisition moves away from all-new, well-bounded, full-up system acquisition program to one of somewhat continuous progressive capability infusion into existing systems regulated by cost.

1.11 T&E and Evolutionary Acquisition

Evolutionary acquisition describes a process by which requirements are met using a phased approach. The requirements can either be well defined up-front for each evolutionary increment, or can be time-phased with firm baselines established at each interval. A common approach today is to infuse technology into existing systems to provide progressive capabilities to the fleet. The emphasis in the mid-1990s shifted from acquiring new systems to that of upgrading existing systems. Evolutionary acquisition moves away from all-new, well-bounded, full-up system acquisition program to one of somewhat continuous progressive capability infusion into existing systems regulated by cost. This reflects the desire to infuse current systems with "state-of-the practice" technology to better manage the life-cycle costs and reap the less well-defined performance gains that may be inherent. Evolutionary acquisition allows the PM greater flexibility when infusing technology into existing systems because there is less need to fully define increment performance up-front. The T&E program can - and should be - tailored to each increment.

An example of an evolutionary acquisition program with well-defined requirements for each increment is the Submarine Acoustic Rapid COTS Insertion (ARCI) effort. In this case, a major upgrade to the submarine's sonar system was developed with an envisioned end-state and well-defined, up-front performance parameters. The program identified three evolutionary phases to meet the required end-state. Each phase addressed a different functional area to be infused with COTS/NDI (Towed array, Sphere/Hull Sonar, and Inboard Processing) with a full at-sea DT and operational assessment made of each increment (see Figure 1-3). Each increment was then released to the fleet on a limited basis for further evaluation while the second phase was developed and tested. This evolution was then applied for the third phase. A final operational test was conducted of all three increments as a system against the established end-state thresholds. The result was that the fleet was able to take advantage of major sonar functional upgrades in

far less time than if they had had to wait for the entire program to be completed. The extended at-sea testing also allowed development and test requirements for each subsequent phase to be refined along the way.

Two examples of rapid COTS infusion for lower life cycle costs and to reduce manning requirements without well-defined performance requirements up-front were the AN/SQS-53D Surface Ship Sonar System COTS infusion and the SMART SHIP Technology Upgrade for Cruisers. The T&E programs for both were tailored to support these non-ACAT efforts. In both cases, the end-state was not well defined at the moment of effort inception, requiring a series of incremental (evolutionary) tests to either further define or correct the direction of the approach. There was a lack of a defined "end state" from which to bound the scope of the test. The requirements were refined along the way, as new information is learned from every DT, OA and OT event.

The upgrade was as well-defined as possible, and measures of the end-state were left as "unthresholded." T&E managers had to oversee the effort closely, as the amount of testing conducted could, at any time, become very open-ended, depending on how big of a subjective sphere of influence was envisioned for the upgrade. OPTEVFOR observed the evolution of the upgrade and its impact on the entire system in the lab, then at-sea. The operational T&E community became involved to the level of complexity of the impact to the system. Not fully defining the performance requirements up-front allowed the system to be assessed in a manner that was more reflective of "testing to learn," versus for "pass-fail". It was found that some upgrades had more value added than others. The T&E program identified the risks and benefits to the sponsor, and it was left to them to decide which aspects of the upgrade to fully introduce to the fleet at what times.

T&E of such evolutionary upgrades is very dependent upon a good working relationship among the test community, and a flexible test plan. The test team must adapt to changing requirements, as performance is characterized as each test progresses. Some tests can be curtailed early, while others would have to be expanded. Figure 1-3 shows how phased DT and OT is used to facilitate getting warfighting capability to the fleet within the framework of an evolutionary acquisition effort.

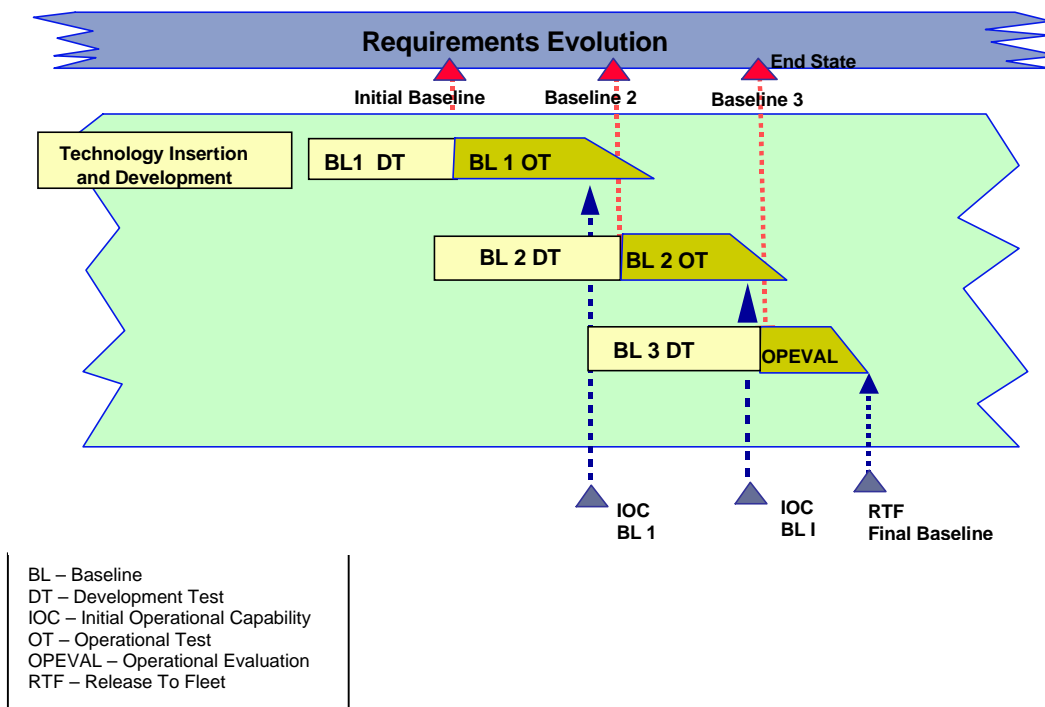


Figure 1-3 T&E Within Evolutionary Acquisition

1.12 Non-Developmental Items

Non Developmental Items (NDI) represent a unique form of acquisition that essentially truncates a system development by utilizing a system, or components of a system, that were already in use by other services or foreign navies. These systems may require some “ruggedization” or modification, an even modified logistics support to meet US Navy requirements. Examples of programs which incorporate these “Non-Developmental” items (NDI) include: the AN/SRN-25 Integrated Navigation System based on commercial equipment, the Versatile Exercise Mine, initially developed for the UK Royal Navy, the Multi-Mission Underwater Remote Operating Vehicle based on remote-controlled vehicles used by the off-shore oil industry, and the Remote Control EOD Tool & Equipment Transporter, which is an adaptation of a ground remote-controlled vehicle in use by police “Bomb Squads.” These programs were aimed at providing the Navy the best available system without necessitating extensive redesign or modification. Often, several different competing NDI systems are taken through DT, with only the best candidate going forward into OT.

1.13 TECHEVAL and OPEVAL

The two key major T&E events of System Development are a system's Technical Evaluation (TECHEVAL) and Operational Evaluation (OPEVAL).

Detailed planning for these T&E events is discussed in Chapter 5. However, they are mentioned here because much of the R&D test program should provide building blocks to these events.

TECHEVAL is usually the final event of Phase System Development DT. It is managed by the program manager and usually conducted by a Navy engineering field activity, with the assistance of the System Contractor. TECHEVAL evaluates the system's technical performance in the Fleet environment, and assesses whether the system is ready for its OPEVAL and subsequent production and deployment. TECHEVAL is conducted partially as a rehearsal for OPEVAL. For shipboard systems, TECHEVAL and OPEVAL are usually conducted on an active Fleet ship.

TECHEVAL validates the technical adequacy of the system as installed and used in its intended environment. OPEVAL provides a snapshot and estimate of the operational effectiveness and operational suitability of the system in actual Fleet use and combat employment. TECHEVAL deals principally with instrumented tests and statistically valid data; OPEVAL deals more with operational realism and the uncertainties of combat. The objective of

OPEVAL is not always to acquire statistically significant data, or a box score of successes and failures (since replications are seldom possible), but rather to gain the most complete understanding possible of the system's capabilities under stress, human/system interaction, Fleet logistic support, and other real world factors. In technical testing, it is generally possible to state the purpose of the test with certainty. In operational testing, the principal value derived is often unplanned, resulting not from the basic purpose of the test, but from realistic aspects that were injected simply because they are likely to exist in actual Fleet/combat employment. Once a TECHEVAL has been completed and the results demonstrate that objectives have been met, the system is formally turned over to OPTEVFOR for the OPEVAL. This process is outlined in Section 5 based upon criteria identified in SECNAVINST 5000.2. OPTEVFOR structures the OPEVAL test plan to prove the system's capability in a realistic operational environment when maintained and operated by sailors, subjected to routine wear—and-tear, and employed in typical combat conditions against a simulated enemy who fights back.

1.14 Unique T&E Requirements

This handbook presents T&E in support of the acquisition process in a manner such that concepts and processes can be applied to a broad range of weapon and combat system developments. Nevertheless, there are unique aspects of T&E that apply to specific system types.

1.14.1 Explosives

Explosives, including propellants, present unique T&E challenges to the test manager. Much of explosive testing will focus on suitability as well as effectiveness (the ability to counter or engage the threat). This includes the ability of the explosive to be safe and reliable in its operational environment. The Navy, unique from the other services, must live and work in its shipboard environment atop its ordnance stores. As such, a tremendous amount of testing and procedures have been put in place to ensure that explosives are operationally suitable. In terms of effectiveness, the approach generally has been to successfully perform serial testing of separate subsystems prior to committing to partial and full-up ordnance tests.

Effectiveness testing includes such items as:

- Target discrimination
- Fuse characterization and performance
- Primary and secondary explosive train operation
- Main charge characterization and performance
- Blast formation and sequencing
- Blast effects
- Bore sight and off axis performance characterization
- Lethality against specific threats

Testing addressing suitability (as outlined in MIL-STD-2106) must address:

- Firing train reliability
- Sensitivity to environment (heat, cold, drop, handling, vibration, shock, altitude)
- Sensitivity to threat stimulus (fire, over-pressure, sympathetic detonation, bullet fragments)
- Storage and aging affects
- Electromagnetic Field sensitivity (HERO tests) – “Hazards of Electromagnetic Radiation to Ordnance”

Any test planning for ordnance must include sufficient time for the serial nature of the ordnance development and test process as well as the full review of test plans, procedures and analysis by safety, engineering and environmental approval boards. All ranges and test sites have safety officers who must be satisfied before testing can be conducted at their sites. Satisfying safety concerns may often require the PM to fund for additional safety precautions, such as unique fire suppression systems, test instrumentation, and physical protection systems to support testing. In addition, most test sites have unique environmental constraints (mammal protection, noise abatement, and airborne emissions) that limit what and when testing can be conducted. It pays for the manager to conduct a thorough evaluation of all potential test sites before proceeding to minimize costs and restrictions. The effectiveness and suitability results of any ordnance testing must first be reviewed by a CNO chartered Weapon System Safety Explosive Review Board (WSESRB) before any developmental or operational test or fielding involving operational forces.

Test parameters for operational testing of explosives supporting the acquisition process are primarily focussed on detonation assurance rather than Probability of Kill (Pk). This is because detonation assurance is uniquely attributed to the warhead whereas Pk usually must take into account delivery aspects of a weapon system. Detonation assurance can be both effectiveness measures such as detection and triggering of the detonator against various targets and suitability measures such as detonation reliability. Explosives, and its packaging, can be considered a weapon system, but a weapon system is typically characterized by a number of sub-systems of which explosives make up one part..

In many cases, because of safety considerations, multiple engineering model configurations are required for development and operational testing. These may include empty warhead shells for physical compatibility and impact tests, instrumented warheads for performance characterization, inert explosive sections for ballistic and weight/moment tests, partially energetic warheads for survivability testing, and full-up warheads for warshot tests. While these only represent a portion of the possible test configurations, the main point is that these various items take resources to develop, certify and manage. These various configuration items and the number that are procured will have a significant impact on the cost and pace, as well as the validity, of testing.

1.14.2 Weapons Systems

Weapon Systems are those ordnance items integrated with other subsystems such as propulsion, guidance, recovery, launchers, loaders, controllers or combinations of each. In addition to all the T&E requirements necessary for ordnance systems outlined above, weapon systems must also undergo T&E to assess the adequacy of these other systems as well as tests of the entire weapon system. These other systems not only include the primary sub-systems listed above, but also the combat control system, navigation, safety and others items usually outside the boundary of the system under development. Not all aspects of a weapon system need to follow a sequential development and test process. Many of the major subsystems can be developed and tested to a degree in parallel. As element maturity increases, they can be integrated with others to form subsystems and full systems that are then tested. The idea is to build a credible story of system risk reduction with these sub-element tests. It is only near the end of development where full-up weapon tests with warheads can be contemplated due to test restrictions and cost. The WESERB is the approval authority for all weapons systems before that can be used in development or operational testing by operational forces. As such, they will review all test plans; procedures and results of major sub-system testing before full-system test are authorized.

Because the major elements of a weapon itself must be compatible, and operate as a system, specific test article configurations must also be developed to support testing. These will include physical mock-ups and empty dummy sections (warheads, boosters, recovery sections, etc.) for compatibility and physical performance characterization. Other tests will require inert sub-systems (warheads, boosters, and squibs) which more closely replicate weight and moment for ballistic and dynamic assessment. While still other tests will require that fully instrumented sub-systems be developed to characterize sub-system to sub-system

performance. These “instrumented sections” are usually very expensive and a key determinant of how much and what type of test data will be collected. A test manager will need to form a “data requirements team” to assess what data will be required and how it is to be collected. These various configurations of test items will require a significant amount of management attention and resources to develop track and certify.

In addition to the weapon itself, its peripherals and other interfacing systems may also require that unique test simulators and stimulators (SIM-STIM) be developed. These SIM-STIM devices replicate the interfaces and query/response characteristics of interfacing equipment. They are necessary if parallel development of peripherals is required. These can include for example, inert missiles for launcher testing and also surrogate launchers used for missile tests. Other examples include combat control simulators for weapon mission data downloads. These SIM-STIM items are a necessity for today’s weapons system development. Their management and engineering is complex and requires a dedicated configuration manager to ensure all SIM-STIM equipment reflects the latest evolved configuration. SIM-STIM equipment will play a major role in at-sea testing as a primary means for training ship’s crew and for troubleshooting equipment problems. In addition, they may be the only source of assets for initial DT given the cost and complexity of actual weapon launches.

Test parameters for operational testing of weapon system supporting the acquisition process are focussed on weapon system effectiveness such as Probability of Kill (Pk), where Pk is ability of the weapon to meet specific success criteria. Examples are the passing below the shadow of a submarine within a prescribed depth as with torpedoes, or within a certain fuse-radius of an aircraft as with missiles. The test manager must always examine the ORD to assess that whatever measures being considered are both measurable and testable. In the case of the MK 50 Torpedo, the success criteria was based on the ability of the weapon to pass with a specific box below the shadow of the submarine to an accuracy of a few feet. Unfortunately, no weapon tracking system at that time could guarantee the desired accuracy. This required a separate parallel development of a new tracking system. Development thresholds usually address sub-system effectiveness such detection range, speed, turning performance,

accuracy placement, fuse detonation, and launch assurance.

Another major consideration for weapon systems testing the conduct of warshot events against surrogate targets. These are tests usually conducted as final demonstrations to verify end-to-end performance of the system including launch, delivery, terminal homing, impact, detonation and blast effects. Because of the unique nature of these tests, they are structured to verify performance already characterized in previous instrumented, inert and other similar testing. Since the weapon and target are damaged or destroyed and instrumentation limited, warshot type tests typically do not yield much new data. Warshot tests require extensive acquisition and preparation of target hulks (environmental clean-up) and documentation for environmental approval. Safety considerations and environmental constraints typically require the test scenario to be staged. In addition a substantial amount of modeling for lethality is required to assess probable fail-safe scenarios. As such, actual warshot tests generally conducted as part of a development effort are rare.

As with ordnance, because of safety considerations, and the cost of live launches, multiple engineering model configurations are required for development and operational testing. These may include inert weapons for physical compatibility and impact tests, instrumented weapons for performance characterization, inert explosive sub-systems for ballistic and weight/moment tests, partially energetic weapons for survivability testing, and full-up weapons for warshot tests. While these only represent a sampling of the possible test configurations, the main point here is that these various items take dedicated resources to develop, certify and manage. These various configuration items and the number that are procured will have a significant impact on the pace of testing.

1.14.3 Sensor Systems

Sensor Systems are generally those functional items which provide the primary input to weapon, combat or Command, Control, Communication, Computers and Intelligence (C4I) systems. Sensors can be individual elements within a system or unique stand-alone systems with their own processing and display items. The later is what is typically envisioned in a stand-alone acquisition program, while the former is usually a subsystem of a weapon or combat system. The primary measure of effectiveness for Sensor systems is their sensitivity to external stimulus and the ability to pass this input to other systems.

Standard measures of performance are:

- Transmit sensitivity
- Receive Sensitivity

- Probability of Detection
- Probability of Classification
- False Rejection Rate

T&E managers must ensure that performance parameters and thresholds selected are testable and measurable. The limitations of the test site and instrumentation capabilities will have a direct impact on the thresholds chosen for controlled land-based testing. Those selected for at-sea testing should contain caveats of the environmental and target characteristics (e.g. target size, intensity, sea-state, humidity, altitude) by which the measures are valid.

The strategy for T&E of sensor systems is comparable to other combat systems, except for one unique aspect. There is rarely a time when all the appropriate environmental or target conditions are present at sea to fully assess the sensor's performance. A full performance assessment can usually only be conducted in a controlled setting such a laboratory chamber or dedicated test site. Such settings allow the test manager to control the environment for a full assessment while also aiding in engineering. Without such control, a majority of test time will be spent trying to identify what stimulus caused the adverse response. Unfortunately for the operational test community, these land-based characterizations are not seen as operationally realistic, though they are understood to be necessary for a full understanding of system performance. What this means for the test manager is that an investment in a land-based test site that allows for stimulus to be varied in such a way that the operational test community can also conduct their assessments will be necessary in most cases. A thorough assessment of all potential test sites is necessary to get an accurate idea of the scope and effort needed for an effective test program. Operational testing at several sites may be required to determine the sensor's ability to operate in different environments.

1.14.4 Combat Control Systems

Combat Control (CC) are systems whose elements link weapon and sensor systems to provide an engagement capability. CC systems generally consist of processors and displays that fuse sensor and command data to provide a tactical picture to the operator along with recommend engagement tactics. Finally, in many cases, CC systems provide guidance and presets to weapon systems.

CC systems must be interoperable with the sensor and command system at the front

end and with the weapon system at the back end. In many cases these associated items are not under the control of the CC program manager, but usually within the same development Command. The scope of T&E of a CC system will be highly dependant upon the breadth and depth of interoperability required. This aspect of CC systems present unique challenges to the T&E manager.

CC systems effectiveness is generally characterized by the ability to place or enable ordnance on target.

Some examples of CC effectiveness measures are:

- Effective Weapon Placement
- Establishing Correct Engagement Criteria
- Accurate Track Generation
- Time to Affect Weapon Presets
- Time to Provide a Go-no-go Decision

Some examples of CC suitability Measures are

- System Hardware and Software Failure Rate
- System Availability When Required
- System Recovery Time

The effectiveness parameters present unique challenges to planning and executing T&E. The ability to affect threat mission abort is only partially related to CC effectiveness, but may be the only measure levied upon the CC system. CC effectiveness is not only a function of the CC system, but also the quality of the sensor input at the front end and ability of the weapon to operate effectively. Thus the test manager must ensure that appropriate caveats are added to any ORD and TEMP CC threshold so that the effectiveness and suitability parameters are testable and measurable.

With regards to test planning, it should include sufficient time for regression and interface tests since there is no such thing as a simple software fix. Even the smallest software change requires extensive, and expensive, regression testing. CC systems require robust land-based test facilities with the capability to emulate the appropriate sensor and command inputs such as acoustic tracks, timing, time of day, navigation, but to name a few. Often, it is difficult to provide the entire spectrum of threats simultaneously in operational testing. Since most CC systems are tested by the same organizations that do the sensor and command systems, the manager can have some confidence that most critical interfaces are accurately defined. Nevertheless, the quality of the CC testing will be highly influenced by the currency and configuration management of these inputs. With regards to the weapon system, critical effort must be made in the weapon system emulation since the results of land based testing will be used as the basis for requesting authority by the Certification Board to engage weapons during the at-sea test phase. The greater the uncertainty with regards to the ability of these

interfacing systems to emulate actual systems, the longer the test and certification process will be.

1.14.5 Command, Control, Computers, Communications, Intelligence, Surveillance and Reconnaissance (C4ISR) Systems

C4ISR encompasses those systems which help define and shape the battle space and enable the use of CC or weapons systems by others within or external to the platform. They include, cooperative engagement systems, surveillance tracking networks, logistics managing systems, battle unit coordination systems, common access intelligence systems but describe a few. C4ISR systems are not usually assessed by their ability to affect mission abort, but rather to enable other warfighting disciplines or to develop situational awareness. C4ISR Systems are some of the most complex development and test efforts, both because of the depth of interconnectivity and because the C4ISR manager rarely owns the interfacing configuration items. These unique aspects of C4I systems present unique challenges to the T&E manager.

With regards to measurement, C4I systems effectiveness can usually only be characterized by the ability to “increase situational awareness”. Specific Critical Technical Parameters are developed to articulate what this “awareness” means in terms of node health and throughput.

Some examples of C4ISR effectiveness measures are:

- Accurate Status Emulation
- Effective Node Sustainment
- Accurate Track Generation
- Time to Accomplish a Task
- Time to Provide a Decision
- Ability to Control a Node or Reconfigurability
- System throughput

Some examples of C4ISR suitability Measures are

- System Hardware and Software Failure Rate
- Network Health
- System Recovery Time
- Node Health
- Man-machine Interface

This wide spectrum of parameters presents unique challenges to planning and executing T&E. The manager should be very conscious about the parameters and thresholds

being proposed for ORDs. High-level measures used for conducting tradeoff analysis may not be testable or measurable criteria suitable for an ORD or a TEMP.

Test planning should include sufficient time for regression and interface tests since there is no such thing as a simple software fix. Even the smallest software change requires extensive, and expensive, regression testing. C4ISR systems generally have the advantage that they can be designed and optimized digitally before committing to procuring expensive engineering models. Nevertheless, a land-based test site, accurately emulating as many of the interfacing systems will be required since computer modeling rarely can address all of the subtleties of the myriad of separately developed peripherals, link systems, controllers and relay systems. C4ISR systems generally consist of processors, data distributors and displays. But because these systems interact with an operator and can process and display more information than an operator can handle, an extensive portion of the engineering and testing will focus on link sustainment, data management and the man-machine interface. The test manager must plan for extensive and adaptable simulation/stimulation (SIM-STIM) as the system matures to aid in addressing this complex and subjective discipline. Despite the rigor of some land-based test sites, extensive at-sea testing and system grooming is required to support the transition from land-based to ship-based testing.

The test strategy for C4ISR systems will involve successive levels of interoperability tests beginning with the assessment of the intra-operability of the directly coupled systems and leading to the assessment of the entire functional stream for a given capability. The second phase of testing will involve the performance assessment between functional streams supporting different warfighting capabilities. The test manager must establish strict configuration control of all test elements and maintain architectural baseline discipline throughout the test process if the program is ever to reach an acceptable end to the effort.

1.14.6 Software Upgrade Programs

Software Upgrade Programs include the introduction of new software to a system baseline to correct previously identified deficiencies, add new warfighting capability or to generally increase the performance of the existing system. Software program upgrades can be conducted on any of the previously listed systems either as a stand-alone program or as an in-service change. This depends upon the scope and impact of the modification. In either case, there is a need to conduct T&E to assess the new performance baseline and assess the impact to the existing system. T&E of these types of programs will usually require assessing the performance of the upgrade at a development facility, followed by extensive testing at a

land-based test site incorporating actual systems. A full series of performance characterization tests will be required and system performance verified before the new software can be introduced into operational units. In keeping with accepted software development practices, each software module is tested and accepted before being integrated. This process is repeated as the system boundary grows. In the case of weapon and combat systems, a full software accreditation review must be conducted by a designated authority. For those system where Communications Security (COMSEC) or Weapons Release Authority is an issue, separate outside reviews must be conducted of the software before testing with interfacing systems. Test managers must plan for these outside reviews which can be quite lengthy depending upon the complexity of the software.

1.14.7 Hull, Mechanical & Electrical (HM&E) Ship Systems

HM&E systems are those shipboard sub-systems that are developed either separately or in concert with the host platform. They are the key systems that support the platform's ability to perform it's warfighting mission. They include items such as the propulsion plant, electrical plant, deck machinery, auxiliary systems, pollution abatement systems, HVAC systems, potable water systems, degaussing systems, turbine generators, chemical/biological filtration systems, firefighting systems, weapons handling gear, etc. Regardless of how they are developed, their common theme is that they must be tested on the host platform, or surrogate, over a long period of time to fully assess their performance. HM&E systems are usually characterized by technical measures such as speed, power, throughput, and capacities, rather than battle effectiveness. As such, the T&E can be focussed to more qualitative assessments. In many instances, HM&E systems are developed and tested by industry to assess performance. In some cases, such as for complex voluminous systems or major power plants, the Navy conducts its own series of tests. These are done to assess the interoperability of the HM&E with other shipboard systems since the prime vendor rarely has a land-based test site comprised of the inner workings of a combatant.

The key issues for the test manager will be how well the land-based site emulates the expected ship environment and the need for long-term dedicated fleet services at sea. In the

first case, ship emulation must not only address performance, physical and electrical compatibility, but also the impact of shipboard environment (rain, wind, ice, vibration) and the variability of services provided by the ship (chilled water temperature, power fluctuations, battle modes, erratic air temperatures, humidity, etc.). These variations in expected conditions must be assessed in controlled settings before HM&E can be considered ready for extensive ship deployment. The second issue deals with long-term performance and suitability assessment in a varied environment. This most appropriate occurs while the HM&E is installed aboard ship. Such extended test period require that ships be specifically designated as test platforms. In most instances the support will be on a “not-to-interfere” basis. That is, the system is to undergo its tests while the ship is busy doing its mission, with limited outside support. Because there is little spare space aboard ships for a duplicate installation, most HM&E replaces it's predecessor for the tests. As such, equipment reliability plays an important factor in keeping the long-term testing ongoing.

1.14.8 Ship Acquisition Programs

If combatant ships were to be procured the way most other Navy systems are, the lead ship of a class would be used as a prototype for the purposes of conducting T&E prior to approving construction of the follow-on ships of the class. However, because of the long time period associated with the design and construction of a ship, it has been agreed that this prototyping approach is not appropriate. Since the majority of the risk associated with a ship acquisition program lies with the new combat systems being introduced to the fleet on that ship, DT&E and OT&E are focused on those systems. Instead of a “full system” TECHEVAL and OPEVAL of the entire lead ship, that T&E conducted on other installations provides the T&E inputs to the production approval decisions for the ships. These installations are: (1) the surrogate platforms used for TECHEVALs and OPEVALs of individual unproven shipboard systems, and (2) the propulsion and combat system land-based test sites frequently constructed for integration of shipboard systems prior to installation in a lead ship. The shipyard's ship construction test and trials phase, as well as the post-delivery test and trials phase, provide additional opportunities that can serve the purpose of DT&E testing and can provide data. Once the lead ship is delivered to the Navy and much of the crew qualification and shakedown period has occurred, OPTEVFOR (if CNO requires that OT&E for the ship) will conduct OT&E to verify the operational effectiveness and suitability of the ship and its ability to perform its missions with emphasis on the newly introduced systems. OT&E of ships also focuses on the integration of these new systems into the ship platform. It is noteworthy that this departure from the norm in the T&E programs of combatant ships is acknowledged in the DoD directives and in Law (Title 10 US Code Section 2400).

The majority of the risk associated with a ship acquisition program lies with the new combat systems.... DT&E and OT&E are focused on those systems.

There is a rare exception to this approach for T&E in combatant ship programs. When a ship design or a combatant craft involves a major technological advance in the hull or propulsion design, the lead ship is designed, constructed, and tested in its entirety as an R&D effort. This is sometimes referred to as a “prototype” ship program. In such a program, the lead ship itself undergoes a full TECHEVAL and OPEVAL prior to the commitment to the follow-on production of ships (i.e., prior to Full-Rate production). In the case of the Landing Craft, Air Cushion (LCAC) program, the decision to authorize the lead production craft was based on OT of the prototype amphibious assault landing craft. FOT&E was later conducted on the production craft to support the full-rate production decision.

Note that previously, the acquisition programs of non-combatant ships, such as fleet oilers, repair tenders, and replenishment ships, did not introduce major new systems and therefore did not require OT&E. However, in recent years, these types of platforms have come under OSD oversight by virtue of being designated ACAT I. This means that a TEMP and some sort of Operational Evaluation is required. The standard practice in the Navy (at least for ACAT II and below ship programs) is that, if the CNO has not required OT&E, the development of a TEMP is not necessary.

1.15 Live-Fire T&E In Ship and Weapons Systems Programs

There is a unique type of T&E in the defense Acquisition, called Live-Fire Test and Evaluation (LFT&E), that encompasses both DT and OT objectives. The Live-Fire legislation, Title 10 US Code Section 2366, addresses testing that must be conducted to assess either the actual weapons effects (in the case of

weapons) or vulnerability (in the case of platforms) against the threat. A program is designated as an LFT&E-designated program if it meets specific criteria as identified in the Law. Once a program is so designated, the PM must ensure that the development T&E effort is structured to conduct sufficient subsystem and system testing to meet the objectives of the LFT&E program. Additional funding and schedule time must be made available to conduct these unique tests and get plans approved by OSD. There are two distinct aspects to LFT&E: the vulnerability of ships and other "covered" systems to enemy fire; and how well our weapons effect mission abort on the threat.

Ships can be considered a "covered system" and their Live-Fire T&E would imply the actual firing of live munitions at these systems. In reality, Live-Fire of ships does not include firing live munitions at a ship configured for combat. There are simply too many variables and unknowns involved (infinite number of shot-lines, burst points, damage scenarios, and combinations of weapons and weapons effects), and limited extra knowledge would be gained at a huge cost. For ships, the intent of the Live-Fire T&E policy is met through a combination of survivability and vulnerability modeling and analyses, ship shock trials, equipment/component shock tests, and surrogate tests. It should be noted that ship survivability features to protect against weapons effects expected to be encountered through the 30- to 40-year life of the ship, are designed into the ship from the earliest stages through the entire program life. The Navy has, on rare occasions, , utilized decommissioned ships as surrogates to test ship's survivability/vulnerability to weapons effects, and these events do involve firing weapons at a decommissioned ship (sometimes tied in with a fleet conducted Sink-Exercise).

LFT&E also addresses weapons "Lethality". The focus here is the assessment of the weapon system's ability to destroy a threat representative target (with main focus on warhead effectiveness). Weapons systems upgrades can be designated a "Live Fire" Program in cases where changes to warhead/fusing significantly change their effectiveness. It must be made clear that LFT&E is not the same as a "warshot" test. The two terms should not be confused. LFT&E is a specific mandated lethality or vulnerability series of tests overseen by OSD. Tests performed during the course of development, utilizing live warheads in rigged scenarios, are considered warshot tests and may be included as part of an LFT&E program.

In both the Lethality and Vulnerability cases, the PM must develop LFT&E plans for OSD approval, and must submit separate periodic reports. Being designated as an LFT&E program is not an easy undertaking – it takes years of planning and a dedicated budget to execute to the satisfaction of OSD. The PM should plan for a dedicated LFT&E manager, under the T&E

manager, familiar with the discipline to meet this requirement.

CHAPTER TWO T&E PROGRAM PLANNING

2.0 Introduction

The T&E planning process is initiated as early as possible in the program, and conducted throughout, to reduce acquisition risks and to provide an estimate of the system's operational effectiveness and operational suitability. For NAVSEA, the major impact of those policies has been to integrate OT&E into the R&D program at strategically significant points. The importance of the Test and Evaluation Master Plan (TEMP) and OT is evidenced in that, the Acquisition Program Baseline (APB), the Operational Requirements Document (ORD), and the TEMP constitute the primary required documents identified in DoD 5000.2.

Chapter One outlined a multitude of methods by which systems are acquired. All new, and modified systems must be operationally tested before the user receives the finished product. The T&E results are the primary vehicle by which the "stakeholders" user are provided the insight into the health of the program and the adequacy of the system. Current DoD procurement practices emphasize reduced cycle times, but that reduction does not necessarily mean less T&E -- it just means smarter T&E. To that end, the recurrent and prominent best practices from past T&E programs point to the continuing need for program and test managers to be personally involved in the planning, conduct, and analysis of major DT&E events. The T&E manager is a vital liaison between the higher echelons, who set the program requirements and approve program continuance, and the engineering team, which actually develops the system. The T&E manager publishes and interprets the system performance requirements from higher management, and analyzes and draws the conclusions of the test reports from below. This chapter describes, based on past experience, the top-level planning concepts necessary to structure an executable T&E program; the extent to which a program manager should direct the test events; and the typical problems likely to be encountered.

DoD procurement practices emphasize reduced cycle times, but that reduction does not necessarily mean less T&E, just smarter T&E.

2.1 Effective T&E Planning

The program manager faces many challenges developing an executable test program. Early tests

should serve as building blocks for subsequent tests. Testing should be structured so that the problems in the higher risk areas are uncovered early, before they jeopardize the success of later testing, when corrections become more expensive.

Programs should avoid transitioning into more complex engineering phases when earlier testing is not yet successfully complete. To do so is to defer problems to a later time, when they will likely be harder to identify and tougher to fix.

Evaluation factors and pass/fail, go/no-go criteria must be established prior to the start of testing. On several occasions, testing at-sea during System Demonstration revealed that a recurring problem, experienced in the Subsystem Development testing years earlier, had not been fully corrected and verified. Programs should avoid proceeding into more complex test phases when earlier testing is not yet successfully complete. To do so is to defer problems to a later time, when they will be harder to identify and tougher to fix.

The program manager needs to be intimately familiar with the planned scope and limitations in each test event. Together with the test team in the Navy labs, engineering activities, contractors, and subcontractors, managers should conduct an early priority/risk analysis and review it periodically to ensure that available test time and other resources continue to be judiciously allocated. The program manager must ensure that test scenarios and detailed test procedures are available to all of the participants prior to the start of testing. The tests must be reviewed, certified for correctness and, where possible, validated through use. This type of review usually uncovers instances where the tests cannot or should not be conducted as planned, and prevents wasted time later because of lost data, lack of proper instrumentation, inadequate attention to training, or loose evaluation criteria.

Effective T&E planning and execution does not happen by accident. T&E is a long-term management discipline that requires good teaming, a thorough understanding of the

requirements, good documentation, and judicious management of resources.

2.2 Managing T&E

There are three critical features necessary for good T&E program management: a well understood management plan, a well-informed test program team, and an effective reporting scheme.

The program manager must coordinate testing conducted by contractors, subcontractors, labs, and Navy field activities. One must maintain a real-time network that provides all participants with the proper information with which to make engineering and programmatic decisions. A written plan can be an effective tool to document this network, to publicize the reporting procedures, and to identify contingency plans and resources. Managers of many larger programs have felt, after completion of R&D testing, that unnecessary and duplicative testing could have been avoided with better planning by the program office. The resultant resources saved could have been more effectively spent in other ways. A management plan, a level of abstraction below the TEMP, can foster the degree of planning that could prevent duplication. Regardless of the size of the program or the number of different organizations, a plan should be developed. While the TEMP is a brief, top-level contract (see Chapter 3) among the SYSCOM, OPTEVFOR and OPNAV that covers primarily the interrelationship of major DT&E and OT&E events, the test management plan orchestrates the details.

2.3 T&E Program Team

Regardless of the size of the program office staff, a point of contact should be assigned to focus appropriate management attention on the T&E program. Experience has shown that the first critical tasks are to identify, assemble, and maintain the necessary T&E engineering support at each of the sites of major program activity. The requirements will obviously vary as the program passes through the different phases. Keeping a good balance of such support, and maintaining a well-defined working relationship, require significant program office attention, frequently more than anticipated. The program office T&E manager needs to ensure early and continuing interface with the OPTEVFOR Operational Test Director (OTD). The OTD should be invited to design reviews, factory and land-based tests, and all DT events. Specific groups that have been used successfully in many programs include TEMP Working Groups (which meet regularly to discuss all related TEMPs and their impact on one another), T&E Coordinating Groups (which meet regularly to discuss T&E progress and to identify issues that need resolution so that tests can be conducted on schedule), and Test Readiness Review/Mission Control Panels (which meet immediately prior to each major test event to examine all T&E aspects of the program to determine readiness to proceed with the event). During production, the need for

flexibility becomes pronounced when the same system is to be installed in different ship types. In one program, poor ship installation so degraded the reliability and availability of the system that it went through several FOT&E exercises before receiving a recommendation for Full Fleet introduction. The resulting demands on the program office and its field representatives were higher than planned and heavily overextended its resources.

An important factor in the efficient conduct of testing is the continuity of personnel and their assigned responsibilities. The same individuals who developed the test scenarios and procedures should be available to support the conduct of the tests. Equally important is that the contractor personnel who are intimately involved in the design of the system and subsystems be available. Experience has shown repeatedly that with poor continuity, testing efficiency is hampered. The team should be well-rounded, composed of knowledgeable personnel, such as:

- A T&E director with significant system level experience and strong leadership abilities to orchestrate the diverse and sometimes conflicting goals of T&E team participants during the schedule-intensive test period, and to make effective on—the-spot decisions that will be supported by the other team members.
- A configuration manager to coordinate thorough and timely reviews and documentation of changes to hardware, computer software, interface drawings, test procedures, and other documentation during the test conduct.
- Test designers with sufficient expertise in their respective areas to identify procedural problems with tests, and revise test procedures on the spot when necessary.
- Test conductors with sufficient experience to recognize T&E problems and to assist in diagnosing them.
- Specialist supporters, such as data reduction/analysis personnel, computer experts, and simulation designers.

The program manager will not usually be in a position to select these personnel, but should be able to recognize problems due to ineffective personnel, and make changes where necessary. But perhaps the more important role of the program manager in this area is to instill in the participants a positive attitude toward using the testing to identify and correct the design errors,

which by the very nature of the program, are fully expected. If measures are not deliberately taken to promote this attitude, experience has shown that some of the participants will be only concerned about “pass-fail” or only in the performance of their individual area, to the detriment of the progress of the overall program.

The program manager should strive to instill in the participants a positive attitude towards proactively using the T&E program to identify and correct the design errors early.

2.4 OPTEVFOR Involvement.

Operational Test and Evaluation Force's (OPTEVFOR) charter allows them to monitor any and all Developmental T&E in each acquisition program. The T&E manager should encourage early and continual involvement by OPTEVFOR for three reasons: (1) Data acquired during DT may be used by OPTEVFOR as part of their OT or as a substitute for it; (2) The OTD will gain a better understanding of the system's capabilities and limitations, thereby enabling him to structure more meaningful and valid OT test plans; and (3) The OTD will be able to provide valuable operational insight that could prove useful in making early design decisions and in preparing for later OT tests. The OTD should be kept apprised of all DT&E planning and results, and invited both to witness tests and to participate in design and other program reviews. During the course of an acquisition program, there will probably be at least three different OTDs as military personnel complete tours at OPTEVFOR and as program reassignments are made because of workload. Each OTD brings to the job his own experience, perspectives, and biases. The program manager must be attuned to this, and anticipate some changes to OPTEVFOR's T&E planning as new OTDs are assigned.

2.5 Requirements Documents

There are many ways that a warfighting capability gets to the Fleet. It can be developed along the guides of the acquisition model; it can be inserted through modification into existing systems as either ACAT or outside the ACAT process; or even procured directly off the shelf. For all these approaches, there is always one governing requirements document issued by the CNO. For ACAT programs, it is the Operational Requirements Document (ORD). For others, it may be a tasking letter from the OPNAV requirements office. In any case, the program manager and the T&E team translate the requirement into “testable” and “measurable” parameters and thresholds. Disputes over requirements, or their definitions, should always be referred to the OPNAV

sponsor for clarification. For Acquisition programs, the system performance requirements reflected in the ORD are translated into testable parameters and thresholds in the TEMP. More will be said about the TEMP in Chapter 3, but the message here is that the ORD and the TEMP require early and continued attention as the most visible expressions of the Navy's performance requirements. Since the TEMP's T&E thresholds are derived from the ORD, attention needs to be given to ensuring that the ORD's performance requirements are testable or measurable, or that they can be translated into measurable criteria in the TEMP.

A frequent problem in planning for a test event is a lack of current, unambiguous performance requirements. The program should use a T&E requirements traceability matrix of thresholds and T&E criteria as a tool to show the traceability of the requirements of the TEMP to the performance requirements at every level of system, subsystem equipment, and component operation. Otherwise, individual design personnel will select T&E points that may not meaningfully contribute in the aggregate to demonstrating total system performance. Systems, both large and small, have been presented for their OPEVAL Readiness Review for which the T&E reports showed little traceability to the TEMP thresholds. Further, the traceability could not be reconstructed, demonstrating that little of it was there to start with.

2.6 Performance Requirements

Performance requirements must be a judicious compromise between what technology has to offer and what the Assistant Secretary of the Navy for Research Development and Acquisition ASN (RDA) and OPNAV feel the Fleet needs and is affordable. Throughout the program, the balance between this famous triad of cost, schedule, and performance is evaluated and adjusted. It is of critical importance, for the purposes of T&E, that performance requirements remain current and visible.

Numerous “requirements” are identified within the Defense acquisition process. These include performance specifications, T&E specifications, Critical Technical Parameters (CTP), Key Performance Parameters (KPP), Measures of Effectiveness (MOE), Measures of Performance (MOP), and Measures of Suitability (MOS). Not all requirements are the same!

A brief definition is provided below:

MOE – Measure of Effectiveness - High-level battle or mission outcome effectiveness (e.g., Probability of Kill)

MOP – Measure of Performance - Mission performance parameter key to meeting an MOE (e.g., Probability of Detection)

MOS – Measure of Suitability - Mission-level outcomes of reliability, and availability needed to execute the mission (e.g., Mean Time Between Mission Critical Failure)

CTP – Critical Technical Parameter - Critical aspects of technical performance, that if not met will result in the inability to meet an MOE/MOS (e.g., Angle of Accuracy, Speed)

KPP – Key Performance Parameter – a critical MOE, MOS, CTP, cost or schedule item that is fundamental to the program, and that can be tracked during development as an indicator of program progress. (e.g., Reliability, Weight, Accuracy)

While each of the above may identify a parameter and threshold, under the context of Acquisition each has a different meaning and use. The ORD will most likely delineate a series of MOE, MOS (e.g., Probability of Kill and Reliability), possibly a few CTP (e.g., speed, acquisition range) and other desired characteristics (e.g., single operator). The technical design agent (TDA) will usually develop the remaining CTP based on engineering studies. The Key Performance Parameters are merely a select subset of these that are identified and reflected in the Acquisition Program Baseline (APB). Only those true “make-or-break” parameters and thresholds for the program should be designated as MOE and MOS.

Program Managers have experienced problems at the Full Rate Production Decision when MOE thresholds that were not met led to an unfavorable OT report. The Milestone Decision point is not the time to attempt to argue that the parameter should have been relegated to a lower level and that missing it is not an indication that the system cannot perform its mission. Only MOE and MOS become the TEMP Part I parameters and thresholds for OT. This will be addressed later in Chapter 3. Other parameters, if critical enough, will become CTPs and the remaining will populate critical aspects of the performance specifications. There is a misconception that everything

in the ORD belongs in the TEMP. Program managers should resist this temptation. Performance requirements should be testable and measurable if they are to be part of the T&E program. Do what only make sense for the program.

Performance requirements, although identified early on, change over the course of development. Tradeoffs will be made in the areas of design, logistic supportability, and affordability as testing proceeds. Modifications to the ORD may be required, and are encouraged, once it has been determined through T&E that the initial requirements were too optimistic or that the cost to achieve the required levels will be prohibitive. *That is the time to modify the ORD!* The concept of balancing performance with that of cost and schedule requirements is called CAIV (Cost as an Independent Variable). The CAIV process identifies those technical parameters that must vary for a fixed cost. For the purposes of Acquisition and T&E, CAIV is that trade-space allowed to the contractor and PM between the threshold and the objective. T&E programs should always first strive for testing to the threshold since that is the minimum acceptable level to which the program is being held.

Program managers of systems that are upgrades of existing in-service systems, of systems that are Navy adaptations of Foreign Weapon Evaluation program systems, or of systems that are available commercially, should take special note. Experience shows a strong tendency in these programs, not only on the part of the program manager, but also the OPNAV sponsor, to devote less attention than necessary to setting performance and suitability requirements, such as logistic supportability. Since a full set of performance thresholds is available from the original system, the tendency is to accept that set with, at best, only minor modifications, and not to update them as development progresses. This has proven many times to jeopardize the program’s continuation when the TECHEVAL does not show the achievement of the long-advertised capabilities.

The Milestone decision point is not the time to be arguing that a particular parameter should have been relegated to lower level and that missing it is not an indication that the system cannot perform its mission.

An example of this is the Submarine TB-29 Towed Array Program, where the ORD mandated an Operational Availability (Ao) threshold reflecting a continuous service system. The Ao concept was a holdover from past experience with submarine sonar systems and the tradition that “all systems had an Ao threshold.” For the TB-29, there was no organizational maintenance for the array itself, other than to come into port and have it replaced. Once the

submarine comes off-station to effect repairs, it is no longer conducting the mission. Clearly the standard Ao definition was not applicable. The array was actually an on-demand system and should have been outlined as such in the ORD. The decision was made to address the issue in the TEMP via numerous caveats to the threshold notes, but in the end it became a nightmare to assess. The lesson learned here is that a major flaw in the ORD cannot - and should not - be "fixed" in the TEMP.

2.7 Number of Test Articles

T&E programs must provide complete and reliable data, used to estimate the military utility of a system. It is therefore essential to budget and fund for a sufficient number of systems and test times to support these objectives. There is a tendency to procure too few test articles and budget for too little integration test time. The result is usually that the extensive series of required performance, reliability, integration, and environmental tests must be conducted sequentially instead of in parallel. The acquisition program should be increased to allow for the collection of statistically valid data, testing, and the additional handling and transportation time for equipment moved from one site to another, which can cause premature equipment aging. Undetected equipment damage may contribute to poor, unrepresentative performance of the system in the later stages of testing.

Some NAVSEA programs have been executed with only a single prototype during development. This always introduces significant schedule risk, and the program often suffers delays as a result. A rule of thumb for non-expendable types of systems is that a minimum of two full systems should be procured for engineering and development: one can be dedicated to performance, reliability, and availability testing; the other can be used for environmental and survivability tests. Once these tests are complete, one of the units becomes the basis for the land-based test site; the other becomes the TECHEVAL/OPEVAL unit installed in the platform.

It frequently proves quite beneficial when there are additional systems hardware, or at least major subsystems, available at the factory or at a land-based test site at the same time one system is installed on a ship for TECHEVAL and OPEVAL. The TECHEVAL and OPEVAL schedules become quite fixed due to the ship's operating schedule and commitments. Access to the equipment during this time is necessarily limited. Another functioning system can serve as a backup for training, and for continued testing and deficiency corrections. On a number of occasions, the "backup" system had to be installed for TECHEVAL/OPEVAL when the primary system was damaged and could not be repaired in time.

For acquisition programs of expendable ordnance (e.g., a missile program), the PM is required to not only find articles for contractor and Navy DT&E, but also to provide the articles for OT&E. It is imperative to

coordinate with OPTEVFOR early to determine their needs. OPTEVFOR, by law, determines independently the number of assets required to conduct their independent OT. It is then up to the PM and OPNAV to assess whether the budget can support this request. If the difference cannot be resolved, the issue goes up the chain, ultimately to the Milestone Decision Authority. Since prototype missiles can cost more than one million dollars each, and aerial targets can cost more than a half a million dollars, it is important to determine the number needed for the T&E program early, both for budgeting and for contract planning purposes. As such, the earlier T&E asset requirements are understood, the better the program is able to identify resources to procure them.

2.8 Efficient T&E Planning

The ability to obtain Fleet services when desired to support a T&E event is difficult. This is especially true for smaller ACAT programs that may not have the high-level support to command dedicated services. A common method of obtaining fleet services is through the "piggybacking" onto other T&E or Fleet training events or by pooling requirements with other T&E programs. For instance, a submarine Combat Control System upgrade program may require torpedoes and missiles, as well as threat target submarines, for a full-scale test. The weapons and sonar systems may also be undergoing some modifications that require T&E. The T&E manager may form a T&E working group composed of other T&E managers to combine T&E requirements and pool resources, to ensure that T&E will be supported by the fleet, and to reduce the total number of assets required. Where possible, DT should be combined with early OT.

Significant cost and schedule savings can be realized through combining phases of DT and OT.

When the OT has a sufficiently operational flavor, OPTEVFOR will frequently agree to combined or concurrent DT/OT. Their agreement is documented in a Memorandum of Understanding (MOU) between OPTEVFOR and the PM, or through inclusion in the TEMP. [The only exception to being able to combine OT with DT is for OPEVAL; TECHEVAL and OPEVAL cannot be combined. By law, at least a portion of the OT supporting the full production decision must be "dedicated," independent OT.] Another area of

potential cost savings is to combine full-scale at-sea T&E with Fleet training exercises or with ship trials and certification tests being conducted in the same time frame. While the latter frequently are not significantly operational in flavor, they can still provide a good source of reliability test data and info from simple scenarios, to augment data to be collected later in formal DT and OT.

T&E managers have also saved resources by obtaining support from organizations outside the Navy active fleet. At times, the Naval Reserve Force and Coast Guard ships have supported DT testing. However, the limitations in operational realism inherent in using such must be taken into account when interpreting T&E results. For example, it would generally not be acceptable to use them for either or both TECHEVAL and OPEVAL.

The T&E manager must carefully monitor the interdependence of related T&E events. Judicious partitioning of the T&E program in a building-block approach is advantageous, because if any part of the program falls behind schedule, it can be detached to a degree to limit jeopardizing the entire program. Another advantage of a well “partitioned” or “phased” T&E approach is that it reduces the scope and schedule demands of the T&E and the data required to be collected and processed per event.

Another way to reduce T&E expenses is by carefully using data from modeling and simulation (M&S) excursions, if validated models are available, to extrapolate data from T&E events and thereby to reduce the scope of full scale T&E.

2.8.1 Keeping T&E Phases Manageable

Partitioning T&E events alone may not necessarily translate into time or resource savings; it just may make it easier to test. Real savings can be achieved by focussing the amount of T&E to the objectives of that event. The number of specific events should focus primarily on exercising the system itself rather than trying to answer entire warfare area engagement questions. It has been shown by modeling that the probability of executing a single complex T&E event is much lower than that for conducting multiple events of a simpler scope. Keeping test phases manageable also allows the T&E program to efficiently respond to unique target or platform opportunities as they arise. The disadvantage of this type of an approach is that more management risk must be assumed to plan for fewer scenarios, shots, and days on-range. T&E that is tied to and combined with on-going fleet exercises benefits from the combined resources of the training and T&E communities.

2.8.2 “Piggyback” T&E Events.

Most lower-level ACAT programs do not warrant dedicated fleet and range services for long

periods of time. For these programs to meet T&E objectives within tight fiscal and schedule constraints, finding a partner to share T&E costs should be a high priority. T&E managers from different programs have sometimes formed “T&E federations” to pool resources and take advantage of unique test opportunities as they arise. Generally, each member keeps a close eye on upcoming events with the fleet training, proficiency, and tactics communities for opportunities to conduct DT. When an opportunity for at-sea testing becomes available, the T&E team can quickly coordinate with all others in the federation to see who can participate and what funding could be used to support it. The T&E federation decides which program would be the lead and what specific objectives are critical, given the unique aspects of the test opportunity. All participants would generally agree to alter their T&E plans somewhat to meet the collective goal.

T&E managers have formed “test federations” to pool resources and to take advantage of unique test opportunities as they arise.

For the Submarine Combat Control System (CCS) MK2 Block Upgrade Program, the TOMAHAWK (TLAM) and Weapons Control System (ATWCS) interoperability T&E portion was structured to match the ongoing TLAM Operational Test Launch proofing events. The missile costs were borne by the TLAM program office and the T&E range costs were shared. The Air National Guard was tapped during some of their regular training exercises to provide one of the many links for the ATWCS part of the test. T&E was supplemented with other rarely used ground and at-sea test platforms, who were very eager to provide low-cost or free support to augment their own, otherwise bland, training evolutions. The tactical and training communities were able to see the latest systems and weapons in action, while the development community was able to stress systems in live, tactically realistic environments at reduced cost.

2.8.3 Harmonizing DT and OT

Combining developmental T&E with operational T&E is a common. Combining DT and OT means that both objectives are met with the same resources at similar times. For instance, for a three-day range event, the first two days could be under the control of the DT

community, with the OTD observing and using the operationally realistic data. The third day could be reserved for the OPTEVFOR to conduct their tests, which the DT community observes. Maintaining a close working relationship between the development T&E team and the operational testers allows T&E to be structured to meet both sets of objectives as test opportunities arise. To accomplish this, there must be some tradeoffs. DT scenarios may have to be structured to add more operationally realistic free play, while OT scenarios may have to be structured with some controlled start-and-stop cues to allow for efficient data collection. Specific test start and end points may also have to be added, to allow more T&E scenarios to take place per day, minimizing range costs. Monthly T&E working group meetings are required to coordinate details and address issues in near real-time.

2.9 Modeling and Simulation

Models and Simulations, collectively known as M&S, are a fundamental part of requirements analysis, design tradeoffs, T&E planning, and actual excursion testing of today's systems. For use in requirements analysis and T&E, the program manager must ensure that the M&S selected has the precision and fidelity appropriate for the role intended. M&S can include engineering computer design models, physically modeled systems or platforms, computer simulations of systems, dynamic simulations of warfare area, or campaign engagements.

Some of the purposes for which M&S can be effective tools for the T&E program are:

- To understand how requirements will be tested and assess sensitivities of the test environment.
- To support pre-test planning, in setting objectives, selecting scenarios, and determining priority parameters to be monitored.
- To identify T&E planning oversights or flawed logic.
- To conduct non-destructive evaluations of scarce items.
- To extrapolate T&E results into other scenarios and levels of force aggregation.
- To represent the input or output of non-available systems, e.g., as a driver or simulator.
- To assess the level of integration required to support the test
- To estimate system reliability growth curves.
- To help identify the conditions under which the MOEs are valid for T&E

When models or simulations are used to support major T&E events, their credibility must be shown. The following are some of the credibility questions to be addressed when considering use of M&S:

- Has it gone through an accreditation process?

- What is the source and the currency of the data against which it was validated?
- What are the major assumptions that were made that, if inaccurate, would greatly influence the results of model or simulation use?
- What are its strengths and weaknesses?
- What field tests have and will be fed back into the model for validation?

The common traps in using M&S that need to be avoided are: (1) unknowingly venturing outside the area of the model that has been validated, introducing unplanned risk and uncertainty; (2) relying solely on heart-of-the-envelope performance data; (3) using specification values instead of actual performance data when the latter is available; (4) continually making worst-case assumptions in an effort to be conservative; and (5) assuming independence between events that actually have some type of dependency or relationship.

DoD policy requires that all M&S used to support operational testing be accredited by the Operational Test Agency (OTA). This means that the OPTEVFOR will conduct their own assessment of to accredit M&S for their use in OT that is planned. Specific policy documents for VV&A used in operational testing can be found at <http://www.cotf.navy.mil/00m&s.htm>. The major criteria for achieving accreditation of M&S for use in conjunction with operational T&E are:

- A complete and accurate description and purpose of the model.
- A summary of the development background and usage.
- A management approach for M&S control, support, and modification.
- Adequate technical documentation.

The OPTEVFOR accreditation process is long (on average 12-18 months) and meticulous, and depends heavily on the accuracy of the documentation used to develop the M&S and the test data available for validation. The PM should plan to conduct VV&A efforts early and in conjunction with OPTEVFOR in order to share the efforts and reduce the costs where possible.

2.10 Interoperability and T&E

Interoperability is a characteristic of system design that allows the accomplishment of tasks using the shared characteristics of two or more systems. Interoperability exists at its lowest level between two components within a sub-system and at a high end at the inter-battle and intra-battle group level. The engineering and T&E

approaches used today reflect the evolution from stand-alone systems in the 60s and 70s, to integrated shipboard systems of the 80s and the inter-battle group systems of the 90s. The emphasis today is to further evolve the T&E process to encompass intra-battle group interoperability. As such, it is a unique occurrence that one system can be tested without input and interface to some other existing system outside the PM's control.

Interoperability transcends two different planes. The first is the interoperability of the system with other direct interfacing systems (e.g., direct hardware interfaces). This is characteristic of systems such as Combat Control, Sonar, and Radar to inter-operate with other ship systems at the ship level. This interoperability has been achieved through the early 1990s with great success, and is relatively easy to characterize, control, and test. The fundamental premise is that for any given functional requirement, an interfacing system can be modified to be interoperable to meet the task requirement.

The second plane of interoperability is more difficult. It is the ability to function with other systems not under the immediate sphere of control and not necessarily directly coupled (hardwired) to the system under test. This interoperability is characteristic of Command Control Communications Computers and Intelligence (C4I) systems, where the ability to be interoperable is dependent upon accepted conventions, architectures, and standards. The fundamental premise here is that all systems must reflect the accepted template. An open system is developed and tested. While the architectural standards may, for this level of interoperability, be defined to a relative degree, the functional aspects of interoperability are much less well-articulated, leading to differences in just how much interoperability T&E is good enough. A well systems-engineered approach to T&E will include interface testing, beginning in the first realm and moving to the second as the development progresses.

Once the system under test has met, in the lab, its own fixed performance parameters, an assessment of the entire system is warranted, using hardware-in-the-loop integration facilities; then in the platform for which it is intended; and ultimately between platforms, if warranted.

The following is an example of a program where its level of interoperability was assumed to be minor, but was in fact found to be quite substantial for the system to be effective. The Launch Expendable Acoustic Device (LEAD) was a surface ship-launched torpedo countermeasure (CM). The measures of effectiveness were solely derived from the CM's ability to defeat the incoming weapon once in the water. The CM was to be launched from an existing launcher used for missile defense decoys upon command from combat control. A

full assessment of the "system" required the launcher to be queued by the ASW detection system, the launch of the LEAD, and finally a corresponding ship evasive action. This type of assessment went well beyond the early DT testing that was performed on the LEAD with emphasis on torpedo interaction. This focus on the interoperability of the ASW combat system, launcher, LEAD, and tactical evasion expanded the scope and duration of testing. Ultimately, the detection-combat control system could not effectively employ the lead without substantial changes to procedures and the design of the launcher. The LEAD device itself was found to be effective, but as integrated into the system, was assessed as not suitable. The lesson here is that interoperability does not end at the program boundary and that T&E must plan for evaluating the system in total.

Battle group level interoperability performance requirements are not yet well articulated, which results in difficulty in determining the how much T&E of interoperability is enough.

The following is an example of a program that met its own requirements and the defined interoperability specifications, but was found to be ineffective when linked to other shipboard combat control systems. The Advanced Combat Direction System (ACDS) program was tested at the land-based test site to its own performance specifications and to the functional requirement; it was found to be ready for at-sea testing. The shipboard installation was completed and certified using an interface check system based on the original ship interface standards. Once out to sea, the ACDS was inundated with so many software trouble faults that it forced the test program to be suspended, ultimately affecting fleet deployment. Apparently, the timing, duration, and types of system messages of the interfacing systems had altered over time, and had become unrecognizable by ACDS. These minor changes, which developed over the life of these systems, were not originally assumed to be of concern for ACDS, but ultimately were. The lesson learned was that a full characterization of interfacing systems, as well as how these systems are used, is necessary to develop an effective land-based test program.

There are some general lessons learned from testing of integrated systems on-

board ships. The following are derived from a C4I upgrade program on a surface combatant:

(1). Understand Requirements - A thorough review must be conducted of all interfacing system requirements up-front and early. The scope and impact of the integration is often underestimated. The ship and combat system architecture assumed in the lab and on paper is not generally what exists in the field. The program must know in advance not only what systems the program is interfacing with, but also how they really are being used.

(2). Understand the Fielded Configuration - A complete and accurate assessment of the installed configuration, down to the workstation level, is critical to understanding how the numerous existing configurations, design changes, and workarounds have been implemented. Managers should consider advanced ships force participation in the design and shipcheck cycle, especially in the case of upgrading a poorly documented existing system.

(3). Dedicated Project Management - A dedicated manager must be in control of a ship-wide integrated project management team. Installation and integration often uncovers issues raised from stove-piped design, installation, and test phases of several different program offices often working at cross-purposes. Poor performance of C4I systems can be directly traced to inadequate system engineering across program office boundaries. A successful program will make use of a dedicated C4I superintendent with accountability for cost, schedule, quality control, and authority over the disparate installation and test contractors.

(4). Coordinated Execution - Planning must ensure that the system installation and test has been coordinated with interfacing system material deliveries, installations, operational commitments, and training events. Performing the test without a true operation of interfacing systems will present an unrealistic picture of performance. Furthermore, schoolhouse support for the system under test must not lag behind the installations. Graduates of the latest schools must be available prior to the installation and test to become familiar with the existing interfacing systems and the new system. Given the lead-time required to stand up a new Navy schoolhouse, onsite training must be provided to the operators. Finally, it is critical that COTS-based systems have sufficient parts on board and have established Navy technical support. Warranty support that only applies in CONUS does not serve ships at sea during a test.

(5). Focused Configuration Management - Given the rapid nature of many C4I installations, a flexible method must be in-place to ensure adequate configuration management of the system, as well as the program interfaces. The lack of adequate design control directly

translates into cost and schedule slips during the installation and test phase.

2.11 Test Site Selection and Constraints

The type of testing required will be a primary consideration in selecting whether testing is conducted in the laboratory, contractor site, field activity, range, or open-ocean. Capability, cost, availability, and constraints will determine where site testing is to be conducted. Since there are numerous government and contractor sites available, it is imperative that the site selected be effective, and that it provide the best mix of realistic operational, environmental, and test measurement conditions to achieve the specific objectives of the test event within fiscal constraints. Environmental regulations concerning protected lands and endangered species legally restrict the timing, type, and scope of testing at most ranges. In response to such constraints, T&E planning must reflect established mitigation measures to remain environmentally compliant. More on this subject is described in 2.20. Such considerations will impact the ability to conduct T&E and the realism of the events planned. For shipboard operations, weather conditions, air and water temperature, sea state conditions, water currents, air traffic density, geographic location, radiation restrictions, and cost are just a few criteria to be considered in site selection. A T&E working group should include those familiar with range costs and schedules as well as environmental constraints and regulations to help develop effective and executable T&E plans. A listing of available DoD ranges and test sites is available at the [Joint Information for Systems Technology, Test and Training \(JIST³\) Site](http://tecnet0.jcte.jcs.mil) <http://tecnet0.jcte.jcs.mil>. Land-based test sites are viable alternatives for some T&E. A discussion of at-sea versus land-based testing is outlined below.

2.12 Land-Based Testing

The emphasis on operational realism in testing has resulted in program managers automatically planning on taking the shipboard system to sea for at least the final portions of DT, and certainly for TECHEVAL and OPEVAL. Sometimes, at-sea tests are conducted even earlier in the program. T&E policies and procedures encourage even more at-sea testing, but this is limited by ship availability and deployment limitations. At-sea testing on fleet ships also limits the program manager's control over testing. Issues such as modifications of the platform, crew training, and other operations related to the T&E program must compete with fleet priorities. Properly conducted land-based testing will supplement and may substitute for

some types of at-sea testing. Properly planned and executed, a land-based test site (LBTS) test program can reduce the cost, risk, scope, and complexity of at-sea tests while maximizing the probability of successful completion of the DT/OT cycle. The AN/SPY-1 Radar is an example of a program that successfully used an LBTS for these purposes. The AEGIS Combat System Engineering Development Site (CSEDS) in Moorestown, NJ is probably the best-known example of a successful LBTS. CSEDS has been used to develop AEGIS upgrades and to train AEGIS cruiser and destroyer crews. So successful and accepted is CSEDS that OPTEVFOR, in a rare deviation to their standard practice of requiring at-sea OPEVALs, conducted the AN/SPY-1B Radar Upgrade OPEVAL there.

Using land-based test sites (LBTS) for testing full-scale prototypes or pre-production systems has become routine for larger systems, but is often overlooked as a means of reducing the at-sea test time for smaller systems. The LBTS can be as small as a single test cell at a contractor's facility, to a large distributed network of multiple sites. The trend is to link all these sites to form a Distributed Engineering Plant (DEP). Linking sites allows the core competency of one site to be tapped as input to another, allowing each site to focus on its unique capabilities, rather than attempting to do all things for all users.

A LBTS is a facility that duplicates, simulates, or stimulates the employment of a system's planned operational installation and utilization for the purposes of conducting DT&E and perhaps some OT&E. LBTSs are often used to test integration of equipment, subsystems, and computer software programs, and is critical to assessing interoperability issues. Test facilities that develop individual equipment, subsystems, and software, or ships and aircraft used as test-beds and general purpose engineering or test facilities, are not generally considered to be an LBTS. LBTSs are occasionally used for the conduct of OT&E (IOT&E) to support LRIP decisions; these typically would involve the larger combat systems or a complete ship's propulsion system. CNO (N091), and DOT&E (for those under oversight) must approve any strategy that involves using OT&E conducted at an LBTS to support a production decision. Such CNO approval is obtained through the TEMP approval process.

With respect to ship acquisition programs, propulsion and combat systems LBTS have occasionally been constructed at Navy Field Activities and Contractor sites to test new equipment, subsystems, and their integration prior to shipboard installation. Technical complexity and risk are major factors in determining the need for such land-based testing. The Navy's propulsion system LBTSs are located at the Naval Surface Warfare Center, Carderock Division (NSWCCD) Ship Systems Engineering Station (NAVSES), Philadelphia, PA.

Typically, they consist of one shaftset of equipment, including main propulsion engines, reduction gear, associated couplings, clutches, associated auxiliaries, and propulsion control systems.

Propulsion system land-based test sites have been used for all surface ship and submarine combatants. Other LBTSs have been constructed at various Government and Contractor-owned sites. The AEGIS CSEDS mentioned above is a well-known example of a combat system LBTS. CSEDS provided a vehicle for conducting numerous DT and OT events on AEGIS combat systems in support of Cruiser and Destroyer programs as they evolved over many years. These combat system LBTSs typically include a combination of equipment, combat system elements (consoles, computers, sensors, and interfaces), and simulators that replicate the systems as installed on the ship.

Several recent major ship programs are using LBTSs in their T&E programs. The LPD 17 ship program is utilizing Contractor Test and Integration Facilities (CTIFs) and Government Test and Integration Facilities (GTIFs) linked in a National Test Network. It is anticipated that in the future, ship programs will likely link with facilities having capabilities that will provide an even higher level of integration/interoperability testing, such as the Distributed Engineering Plant (DEP). In the case of the Virginia Class Submarine program, the ship's combat systems' electronics will be tested at the Command and Control System Module Off-Hull Assembly and Test Site (COATS) facility located at the shipbuilder's shipyard. COATS will provide the opportunity to perform a large share of the systems testing prior to ship Float-Off.

Major benefits from using propulsion and combat system LBTSs in ship programs are: (a) early identification and resolution of hardware interface and integration problems (prior to shipboard installation); (b) early demonstration of complex hardware/software integration; (c) early identification and resolution of computer software problems; and (d) test-bed for conduct of IOT&E to support LRIP decisions for ship programs.

The utility of land-based testing is determined primarily by how closely the actual operating environment can be approximated during the tests. In the Class B-1 Radar program, although effects of pitch and roll on radar performance could not be assessed at a land-based site, placement of that site on the shore enabled detection of real targets in the presence

of close-in sea clutter. In one rare instance, the land-based detection performance data was used when abnormal propagation conditions invalidated the at-sea test data. Program schedule constraints precluded an at-sea retest, and the existence of LBTS data saved the program from failure. Another factor affecting the utility of LBTS is the availability at the site of those systems with which the system under test must interface.

Land-based testing does not always require a specialized facility. In many cases, Navy laboratories or field activities can serve as excellent T&E sites. Shore-based engineering centers can sometimes provide an inexpensive and operationally realistic environment without the scheduling and installation constraints of a fleet asset. In general, the LBTS environment allows for better control of the equipment under test, the test variables, test personnel, logistics, and communications. Significant advantages of a LBTS include:

- Highly controlled and repeatable T&E opportunities (important when modifications demand retest).
- Early discovery of system operational limitations/problems.
- A wide range of variable permutations can be investigated without tying up scarce (and expensive) platforms and other critical assets.
- "Early life" system failures occur and are quickly corrected on land rather than at-sea (where they can unduly delay or force cancellation of tests).
- Some system capabilities can be fully stressed only at LBTS. (Consider the cost of several hundred aircraft to test maximum track capacity of an automated tracking system, versus simulation at LBTS.)
- T&E team knowledge of the system is increased prior to at-sea T&E, resulting in better at-sea testing.
- The personnel and facilities required to correct problems are more readily available, and required fixes can be implemented more rapidly.
- Availability of shore-site technical and clerical help and assets permits on-site revision of T&E plans/documents as required by initial results, and offers an opportunity to begin drafting preliminary results.
- Opportunity for training of ship's force personnel.
- Opportunity to begin early development of operational doctrine/guidance, which can be used to minimize the impact of system problems that cannot be solved through modification.
- Allows OPTEVFOR's early involvement through combined DT/OT testing and their monitoring of the DT at the LBTS.
- Assures the program manager that the system has attained a reasonable level of maturity in design, performance, and reliability before he commits his resources to a high-risk, at-sea exercise.

- Allows validation of software builds before use on TECHEVAL/OPEVAL system, as well as on eventual permanent shipboard installation.
- Provides an early opportunity for validation of the system's maintenance concept.
- Allows validation of installation checkout procedures for use on the TECHEVAL/OPEVAL ship, as well as on eventual permanent shipboard installations.

2.13 LBTS Role In Interoperability

An LBTS, whether at a contractor, laboratory, or a field activity site, is usually optimized at the sub-system level. Some of the more complex combat system evaluation and test sites, such as for AEGIS Combat Systems Engineering Development Site (CSSEDS) in Moorestown, N.J., are designed to address performance at the entire air defense system level. Nevertheless, as systems become increasingly interdependent and the emphasis on battle-group interoperability increases, so must the LBTS address this interdependency with the use of multiple test sites. The idea behind linking LBTSs is to capitalize on one site's unique core capability to test a system at another site. For example, if the Anti-Air Warfare (AAW) defense system is designed to be queued by the Electronic Warfare (EW) system, this interoperability must be tested. Rather than trying to sub-optimize the T&E by replicating the EW system at the AAW test site, the AAW site is linked to the EW LBTS. The linkage of sites provides a distributed web of engineering and test capabilities available to the program manager. Within the NAVSEA community, an example is the Distributed Engineering Plant (DEP). The DEP is the functional linking of specific sites, protocols, and operating procedures available for Battle Group interoperability T&E, in particular. A full description of the DEP is available at the Naval Surface Warfare Center, Dahlgren site at <http://www.nswc.navy.mil/dep>.

2.14 Specialized T&E Resources

T&E of today's ships and combat systems at times requires the use of specialized T&E resources to conduct testing. These can include advanced threat representative targets, unique range upgrades and systems, and up-to-date threat simulators/emulators. Those that act as surrogates for threat assets, such as modified torpedoes and missiles, or platforms such as the Self Defense Test Ship, USS DOLPHIN Test Ship (AGSS-555), and EX-SALMON SSK Bottom Target must be acquired, modified or scheduled years in advance. The cost of developing, modifying and using these new or upgraded T&E

assets, are usually borne by the PM. In some cases, if the asset can be used by other programs, then some institutional funding can be found through the CNO sponsor or N091 to help support the effort. The T&E manager must define the requirements for such assets as part of the overall planning process early in the program, in consultation with the DT and OT T&E community, so that their long-term needs can be planned and programmed. In some cases, Threat/Target Validation IPTs have been established to assess the types of threat surrogates required to support T&E. Specific details concerning each of these assets can be found at Navy Warfare Center or Field Activity websites. The NAVSEA T&E Office participates in these forums. A good source for such resources can be found at the [Joint Information for Systems Technology, Test and Training \(JIST³\) Site http://tecnet0.jcte.jcs.mil](http://tecnet0.jcte.jcs.mil)

2.15 Threat Target Validation

DoD policy requires that all targets used as surrogates to represent the threat during T&E events undergo a validation process. For those programs under OSD oversight, this validation report is submitted, and is approved by the Director, Operational Test & Evaluation (DOT&E). This Threat Target Validation (TTV) process essentially examines available and planned surrogates, and compares characteristics critical to resolving MOEs with the threat. TTV efforts must be made early during program design to ensure that resources are planned for surrogates which may have to be developed or modified to support future T&E. Figure 2-1 outlines a TTV process flow for undersea warfare surrogates used in T&E within NAVSEA/PEOs.

2.16 Designing for Testability

The ability of a system to be tested effectively and efficiently, during production and while in service, must be designed into a system EARLY. This requires a strong testability mentality and commitment among the entire program team. When the needs related to such testing are fully recognized from the beginning of the design process, they are adequately accommodated. During early design, the performance characteristics to be measured during production and in-service testing must be identified. Data on these characteristics must be collected during development, and ready access to those parameters must be designed. The design also must enable rapid and accurate assessment of the status to the lowest repairable element when deployed. An easily and completely testable design that can be inspected without disassembly, adjustments, special environmental conditioning, or external equipment or stimuli for monitoring of responses, is amenable to economic production. Testability is also defined as the ability of a design to facilitate T&E. There are many times when unique data extraction requirements are needed in near real-time or real-time applications. This is especially true for systems

that may not be recoverable after test, such as missiles, UAVs, and UUVs. These T&E enablers must be engineered into the system early. A good vehicle for getting these issues identified is through early T&E Integrated Process Teams (IPT) with feedback to the engineering IPT. A frequent misconception is that ingenuity in the design of manufacturing test equipment can compensate for deficiencies in the testability of prime hardware. In reality, not much can be done to "add on" testability if provisions for it were not made in the original design. Spend a little money early in the program to design in testability. No amount of expensive breakout boxes, extender cards, or maintenance assist modules (MAM) can compensate for a poor testability design. A well-designed BIT also lowers system mean—time—to-repair (MTTR) values (which helps the program achieve the required Ao) and reduces operations costs throughout the life of the system. A fully maintainable design is quite likely to be highly producible.

2.17 Data Collection and Analysis

A major cost and schedule driver in any test program is collecting, analyzing, and reporting data. The test manager should have a prioritized list of objectives prior to the start of any test, with contingencies addressing surprises. All participants must be aware of their responsibilities for test operations, data collection, problem resolution, data analysis, and test reporting. An important lesson learned in data collection has been to provide for the recording of all pertinent data, and then to be selective when determining the data that are to be reduced for further analysis. A clever technique sometimes used to verify adequate planning for data recording and analysis is to have the T&E report outlined before the T&E is conducted, and work backwards to determine data needs.

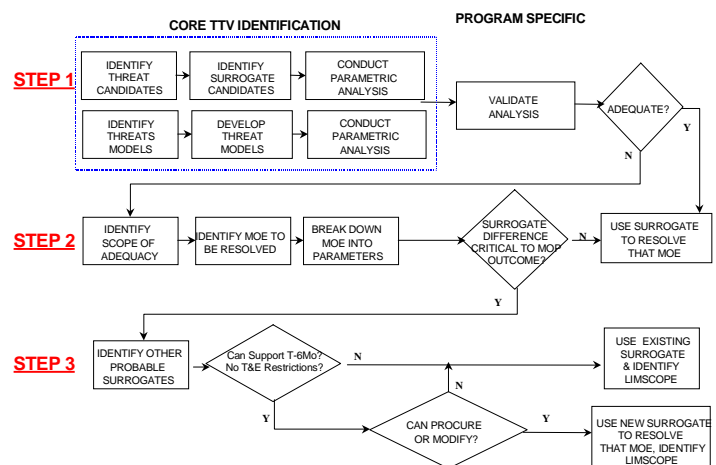


Figure 2-1 Threat-Target Validation Process

There is nothing more embarrassing (and costly) than completing a test, only to find, during the writing of the final report, that some critical data was overlooked, and therefore some needed conclusions cannot be drawn or substantiated. With regard to the resolution of problems uncovered during testing, procedures must be defined in advance of testing to determine which problems will be recorded, how they will be described, and who will have follow-up action to resolve them. On-site expertise should be available so that only a minimal number of problems need to be referred to other locations for resolution or approval. Prior arrangements should be made with other organizations to ensure that the resources, priorities, and working relationships are established for expeditious handling of, and attention to, the problems that occur during testing.

A thorough data requirements analysis must be done early to assess the type and amount of data to be collected during each phase. The approach should be to collect data with high fidelity where it is easiest to do so. The thought is not to drive testing with data requirements, but to drive data collection with the testing planned. Since fleet exercises and tactical proficiency events are the source of the majority of at-sea opportunities, data requirements must take advantage of the existing fleet performance logs and collection procedures. It is better to obtain more data of lower fidelity, than to obtain almost none of the data more difficult for operators to collect. The data collection scheme must also be designed to allow at-sea T&E data to be supplemented with other contractor certification test results, to further increase the database available for inclusion in any final assessment. Considerations for how best to design systems for efficient data collection are found in Section 2.13, "Designing for Testability."

For data analysis, a rule of thumb is to complete sufficient analysis to reach an 80 percent confidence level, allowing for defensible conclusions and recommendations. . Nevertheless, it must be remembered that a test event is merely a snapshot of performance at a specific time. To infer total system performance over the life of a program with only a snapshot is misleading, no matter how statistically confident the data set. In many cases, there are not enough resources to obtain statistically significant sample sizes or for high confidence levels. In such cases, the T&E reports and analysis must clearly identify the degree of confidence that can be inferred from the data.

2.18 Environmental Qualification Testing (EQT)

EQT is conducted during system development to assess the ability to withstand environmental extremes in controlled tests. These extremes, such as high-impact shock and temperature, are unlikely to be experienced

during a DT or OT event. The test unit(s) used for EQT are generally not used for actual TECHEVAL or OPEVAL, due to the risk that that component may be stressed or degraded, leading to failure during these events. The PM should plan for a dedicated pre-production unit to be used for EQT. The following is a listing of the most important environmental characteristics in the T&E programs for shipboard systems, along with the current specification that identifies the T&E requirements:

Parameter

Military Environments	MIL-STD-810
Shock testing	MIL-S-901
Electromagnetic compatibility (design and test)	MIL-STD-461 and 462
Vibration	MIL-STD-167
System Safety Program	MIL-STD-882

A system should have successfully completed environmental T&E before entering TECHEVAL. A full characterization of system performance is needed under all expected operating conditions. This is especially true for COTS-infused systems, where the degree of robustness may be unknown. Similar rationale applies to the other environmental qualification areas. For example, a processor installed and designed to work in a protected area inside a commercial aircraft will not necessarily be able to withstand the stresses of the shipboard environment (salt spray, humidity, temperature, vibration, EMI, etc.), not to mention the possibility of very different signal and functional interfaces. The PM needs reliable information (vendor test data, for example) or the PM's own independent T&E to verify system ruggedness.

2.19 RM&A

Reliability, Maintainability and Availability (RM&A) are so significant to the Navy that it is very rare that they are not included among the key performance parameters for a new system. The requirements for developing, estimating, and designing-in RM&A requirements are well addressed in other publications, (search under "reliability" at the **DoD Deskbook** site at <http://www.deskbook.osd.mil>), so they will not be treated in depth here. However, it is important to recognize that the quantitative RM&A thresholds must be engineered and reaffirmed throughout the program. They are very dependent on system design, configuration, failure definition/criteria, intended operating profile and duty cycle, spare parts loading, training, and all of the integrated logistics support elements. Changes in any of these factors during development may impact the ability to achieve RM&A thresholds.

In DT testing prior to TECHEVAL, at least one maintainability demonstration should be performed in a controlled environment to assess the design of the system. Another maintainability demonstration should be conducted during TECHEVAL. It should be structured very similarly to fleet operations to assess training, repair kits, and the system's capability under stress. With one controlled maintainability demonstration and one real-world demonstration completed, the results can be evaluated prior to entering OPEVAL. Where possible, agreements should be concluded in advance, so that OPTEVFOR can observe DT events that include the "real-world" maintainability demos and other R&MA testing on the system. If this DT data is usable by OPTEVFOR to supplement OPEVAL results, some reduction of OT test time can be realized. Also, in TECHEVAL, the later maintainability demonstrations can provide additional opportunity for crew training prior to OPEVAL.

An important criterion that should be identified when specifying RM&A parameters and thresholds is how failures will be counted. It is important that OPTEVFOR and the PM have general agreement as to how failures will be counted in OT and DT. Generally, critical operational mission (mission preventing) failures are counted in MOS reliability and availability calculations. Minor failures are generally counted toward Critical Technical Parameters, if so desired. Software-related failures are counted the same as hardware failures, so allowances for software failures in the reliability threshold must be made if a separate software reliability threshold is not specified. The types of deficiencies, defined as critical, major, and minor failures, must be identified for each system's operating profile. With regard to operational availability, the mission profile upon which a given operational availability (A_o) is based should be included or at least referenced in the TEMP. This profile should include a time-phased description of the events and environments the system will experience during a specific mission and the portion of total time the system is active in its various operating modes and the duty cycle.

The PM must have a clear understanding of the interrelationship of RM&A. Clear guidance in the definition is found in the DoD Operational Suitability Guide; however, how these definitions apply to the system is critical to selecting the right conditions in the TEMP, which makes the parameters and thresholds valid. First and foremost is a clear delineation of which systems, subsystems, and components are covered under the definitions. Remember, some peripherals, such as controllers, printers, and storage devices may fail, but may not be mission-critical, and thus would not be counted under Operational Mission Critical Failures or Mission Availability. Additionally, the definition for reliability and availability may change for self re-configurable systems, or those which operate in some

acceptable degraded mode of performance. Continuous System Operational Availability (A_o) implies the ability to fix the failure at the operator "O" level. It is important to caveat subsystems that cannot be repaired at the "O" level, where the test platform must come off-station to conduct repairs, effectively ending the test (mission). Where necessary, a referenced RM&A description document may be required for large systems to help testers adjudicate issues for scoring.

The late 1980s marked a period of transition for measuring RM&A with regard to hardware and software thresholds. During this period, software became a significant facet of the design – to the point where it impacted operational suitability in ways that hardware never did. The standard RM&A tools used were not designed to address software faults, failures, and system restarts. The DoD recognized this by establishing separate hardware and software reliability and maintainability thresholds in the ORDs. Software development and testing evolved from an art to a disciplined engineering process in the 1990s, as did the understanding of software reliability and maintainability. The acquisition process then shifted to one threshold to cover both hardware and software failures. By the end of the 1990s, software complexity again became the dominant reliability driver in most software-intensive systems. The reality was that hardware was not failing, as it had in the 1970s. The trend now has reversed again, with software thresholds being reflected in the ORDs and TEMPs in addition to those for hardware.

2.20 Computer Software Testing

Software is a fundamental functional element of any major shipboard system, and offers a significant challenge to T&E planning. Compared to hardware, there are several aspects of computer program testing that make it challenging: the need to establish meaningful system operational requirements; to establish adequate test specifications (testable requirements, effective pass/fail criteria, and error definitions); to clearly define all affected interfaces; and to enforce strong configuration management.

The need for efficient T&E planning is more critical for software than it is for hardware because of the great expense involved in development, and the need that the software be proven adequately operable prior to system integration.

Efficient T&E planning is more critical for software than for hardware, due to the great expense involved in development, as well as the requirement for tested and certified software elements prior to integration. IEEE STD 12207 Addresses system software development and establishes the requirements for developing and testing computer programs/computer upgrades identified in the Software Development Plan, and the systems in which computers are embedded. These requirements will not be described any further in this manual. However, it must be emphasized that all of NAVSEA's experience points to the need to follow the spirit and intent of these publications to successfully demonstrate system capabilities. While there are numerous ways to measure software development, the following six Software Development Metrics, derived from the DoD's Program Manager's Guide To Software Acquisition Best Practices ([URL: www.spmn.com/best_practices.html](http://www.spmn.com/best_practices.html)), are good T&E measures that track the health and adequacy of software intensive programs over time:

- Requirements Traceability – Percent of mission requirements traced to code and test.
- Fault Profiles – Priority, Category, and Status.
- Requirements Stability – Number of ECPs written against software requirements.
- Design Stability - Percent of changes in software design.
- Breadth of Testing – Number of requirements tested over the number of test requirements.
- Depth of Testing – Number of paths/decision points exercised at least once over the total paths/decision points.

Additional guidance can be found in the DoD Program Manager's Guide To Software Acquisition Best Practices. SECNAVINST 5000.2 also gives high visibility to the need for system's software maturity to be demonstrated prior to commencing OPEVAL. In fact, the program can not proceed into OPEVAL with any outstanding, unresolved Priority 1 or 2 software faults or (Software Trouble Reports (STRs)), otherwise a waiver must be granted from CNO (N091).

2.21 NEPA and Environmental Compliance

In many locations, the ability to conduct T&E is directly impacted by the environment and Federal restrictions outlined in the National Environmental Policy Act (NEPA). T&E events, such as high-impact shock tests, have a potentially significant impact on the environment. OPNAVINST 5090.1 outlines the processes and procedures required for assessing and reporting the potential environmental risks of a development program, and the constraints under which tests can be conducted. This instruction requires the PM, as the "Action Proponent," to conduct a Programmatic Environmental, Safety and Health Assessment (PESHA)

of the entire program. A portion of this assessment will cover the T&E program. Currently, there are no exclusions in OPNAVINST 5090.1 specifically for T&E events. Therefore, PMs must judiciously choose sites, scenarios, and times where T&E can be conducted to limit the need to generate extensive and expensive Environmental Assessments (EA). The Navy requires that a Finding Of No Significant Impact (FONSI) be obtained from CNO (N45) prior to the conduct of any test event where there is potential to affect the environment. It is strongly recommended that early communications with Field Activity and Range representatives be made to select sites, times, and scenarios that already conform to existing, environmentally approved procedures.

PMs can avoid need to generate extensive and expensive Environmental Assessments by being judicious in selecting the sites, scenarios and times of major T&E events.

2.22 Operations and Data Security

Operations Security (OPSEC) is defined as "the process of denying adversaries information about friendly capabilities and intentions by identifying, controlling, and protecting indicators associated with planning and conducting military operations and other activities." Every T&E program must consider OPSEC. The Program Protection Plan (PPP), for which the program manager is responsible, is the document that delineates the planning, coverage, and responsibilities for any T&E program requiring OPSEC and containing classified data. Planning needs to begin early in program development and continue formulation as the program progresses. Elements to consider with regard to OPSEC include:

- OPSEC is not just a plan, it is a philosophy. It will reflect how you will conduct secure testing.
- Proper security equipment takes time to procure, integrate, test, and learn to operate.
- OPSEC guidelines and rules for the program must be published early. This gives everyone involved a clear idea of what can and cannot be done during testing, and thereby reduces some "false starts" in T&E planning.

- Avoid combining T&E data with ship operations data when possible. T&E data has sometimes been confiscated by the ship's force for this very reason. Keep in mind who will have control of data links and the T&E data throughout testing when preparing OPSEC guidelines and rules.

Along with OPSEC, program managers must consider how any sensitive T&E data will be protected during system development. This is equal to providing measures that prevent such data from being compromised. The PPP should be approved and implemented prior to any testing where classified data will be gathered. When developing this plan, program managers should be aware of the numerous secure capabilities available at RDT&E activities before developing a unique, program-specific capability. Details and guidance on the preparation of a PPP can be found in the Defense Acquisition Desk Book.

2.23 T&E Funding

In general, the program manager must budget and fund the expenses of the T&E program, including OT&E. Examples of these are the costs of: test articles, expendable ordnance, laboratory and contract analytic support and instrumentation, data collection and reduction, training for system operation and maintenance during T&E, maintenance and logistic support costs, and use of Navy ranges. OPTEVFOR usually requests that the program manager provide all funds to support OT&E to them, rather than directly to the organization that will actually expend the funds, but this is sometimes negotiable. The PM has the right to question all funding requests by T&E participants, including OPTEVFOR. A significant cost driver in OT&E is the analysis of collected data. OPTEVFOR must be involved in discussions concerning the responsibility and the amount of resources available to support various assessments.

In general, the program manager must budget and fund the expenses of the T&E program, including OT&E.

However, the following costs are generally not borne by the program because they are budgeted and funded from other sources: Fleet travel for training, operating cost (including fuel) for active Fleet ships and aircraft, targets, OPTEVFOR salaries, travel, and non—program-related costs. However, the program may be asked to supplement the funding of these items on a case-by-case basis.

2.24 Fleet Services and Support

For the at-sea testing of shipboard systems, Fleet support is critical and requires long lead-time to

schedule and confirm. Because of heavy operational commitments, Fleet services are scarce and not always available at the time or on the level preferred. Timely planning and scheduling are essential to ensuring their availability.

Fleet services are divided into three levels. They are dedicated support, concurrent support, and not-to-interfere basis (NIB) support. Dedicated support is provided when the participating ship(s) and aircraft are precluded from other missions; they are dedicated to the T&E event during the time scheduled. Concurrent support means the assigned ship(s) and aircraft may need to participate in or be impacted by other missions. NIB support means the services are available to support the T&E requirement, but those services must not significantly impact on other assigned missions. The level of support must be based on the type of testing and other factors, such as Fleet availability, range requirements and availability, electromagnetic emissions restrictions, and test duration. Dedicated (Priority 1) services are much more difficult to receive than NIB (Priority 3) services; as such, contingency plans should always be ready. The level of support provided is based on the requirement, not on the ACAT. Even after careful planning, meticulous documentation, and timely submissions, the services provided may not be exactly what is wanted when desired. Even after services are scheduled, events can delay or cancel out the services. Therefore, in planning tests requiring Fleet services, the program manager should plan for scheduling and funding contingencies to minimize the impact of such problems.

About six months prior to a given quarter, CNO N912 solicits requests for ship and aircraft services needed for that quarter. The NAVSEA T&E Office, in turn, distributes this letter within NAVSEA and its associated PEOs for the specific needs for that quarter. The response to the NAVSEA T&E Office is usually due six weeks later, after which the T&E Office provides a consolidated response to OPNAV. Soon thereafter, OPTEVFOR, as N091's agent, participates in the Quarterly Fleet Scheduling Conferences conducted by the Fleet Commanders. After the conferences, OPTEVFOR publishes the results as they pertain to the T&E support requests.

Several aspects of this process are worth special notice:

- NAVSEA submits the request for Fleet services to support DT and Combined DT/OT events; inputs for dedicated OT are submitted by OPTEVFOR.

- The requirements for services that encompass more than one quarter must be resubmitted each quarter.

- OPNAV does not usually provide Fleet support services to an acquisition program without an approved ORD or TEMP. If a program manager does not have an approved TEMP, there are certain exceptional circumstances in which the NAVSEA T&E Office can assist in obtaining the services. If the program is a non-acquisition program (i.e., technology development only or if the program is a Foreign Weapon Evaluation (FWE)), a TEMP is not required, but Fleet services can be authorized by obtaining from OPNAV a T&E Identification Number (TEIN). In the case of such non-acquisition programs, the number is given a "K" prefix for technology development programs or an "F" prefix for FWE programs.

- Program managers must not go directly to Fleet Commands to formally request such services. Every time a program manager sends a message directly to the Fleet requesting services not previously arranged, he jeopardizes other scheduled services. All requests, even urgent ones, must be directed to the CNO N912. Urgent requests, also referred to as "Emergent Requirements", must be requested by message to CNO N912.

- OPNAV prioritizes the Fleet service requests for each quarter. The highest priority is a "CNO Priority ONE." If obtaining the Fleet services in a given quarter is critical to the execution of the program (i.e., the testing cannot be delayed), a program manager can improve his chances of getting a CNO Priority ONE assigned. This assignment can be obtained by having the OPNAV Program Sponsor (Flag Officer level) send a memorandum requesting such a priority, with proper justification, to N091. The OPNAV deadline is usually one week after the NAVSEA T&E Office due date for Fleet service inputs. Priority ONEs must be justified each quarter by the sponsor and will not automatically carry over. The priority applies only to the indicated Fleet support, not to the entire R&D project. It should be realized that this priority is what OPNAV assigned the services, relative to other requests for services to support T&E. Generally, Fleet Command will further prioritize support of each T&E exercise among the other tasks they have for their ships and aircraft that quarter. Thus, a CNO Priority ONE is still not a guarantee of obtaining the needed services. Further improvement in obtaining critical fleet services can be achieved by having a knowledgeable program office representative attend the fleet scheduling conference.

- In addition to the approval of OPNAV and the Fleet Commander, the program manager must

coordinate closely with the cognizant NAVSEA Ship Program Manager in arranging for shipboard installation of a system. Although the planned installation is temporary, it may impact the ship's weight and moment, safety, and logistics support.

2.25 Supporting Milestones and Decision Points

Planning for the decision milestones and top-level planning for the T&E program are interrelated and interdependent. One of the major inputs at the milestone review are the results of T&E to date; conversely, the T&E program must be structured to provide meaningful and substantive T&E data at the decision points. The timing of milestones and the scope of and type of testing accomplished will vary from program to program. The T&E must be tailored to the program's risks. However, by the Full-Rate Production Decision, the T&E program must demonstrate that the system is technically sound, operationally effective, and operationally suitable. The most important thing to remember in program planning, as well as in detailed scheduling, is to build in slack to anticipate a reasonable number of unforeseen problems and to allow for the correction of problems uncovered during testing. Experience has shown that, for a variety of reasons, we tend to be too optimistic in our planning and tend to be more schedule-driven than performance-driven. We frequently approach the Full-Rate Production decision with some significant unresolved technical problems. When the start of production must be deferred for this reason, it can provide visible "bad press" for the program and can even jeopardize its continuance.

Nearly all programs have a decision point for low-rate initial production (LRIP) and for full-rate production (FRP). Per Title 10, United States Code (USC) 2399 and the DoD 5000.2 Instruction, there are three reasons that LRIP units are approved:

- To provide production configured or representative articles for operational tests
- To establish an initial production base for the system
- To permit an orderly increase in the production rate for the system upon successful completion of operational testing.

Only the first reason is directly applicable to executing a test program. There may be other business-related reasons why production begins while still in development. In any case, the Milestone Decision Authority (MDA) must be provided with an indicator of program risk

based on T&E to date. The LRIP decision allows production to begin, although the system will not have demonstrated all of the capabilities required for Full Rate Production.

There is no set amount of T&E, either DT or OT, required per the SECNAVINST 5000.2 to support an LRIP decision point. Nevertheless, an MDA would benefit greatly from (and reduce exposure to criticism) having OPTEVFOR make a preliminary assessment of the system to support the decision. The objective is to

use T&E on the immature system to point out high-risk areas to the MDA. The primary purpose of the TEMP is to document the agreements with OPTEVFOR over the scope of T&E and the level of system maturity to be expected at the LRIP and full production decision points.

CHAPTER THREE DOCUMENTING T&E

3.0 Introduction

The necessity of good T&E documentation cannot be overstated. The amount of documentation necessary can seem overwhelming, but it is critical to supporting the engineering and programmatic decisions involved in the acquisition process. Good T&E documentation captures direction to the testers and evaluators, prioritized test requirements, sound resource allocation, and a reflection of design maturity and risk. Of all the documents generated to support effective T&E planning, such as those required for describing high-level strategy, test phase and event conduct, evaluation and assessment, reporting, and summarizing DT, the Test and Evaluation Master Plan (TEMP) and the Live-Fire Test and Evaluation (LFT&E) Plan are required by statute.

3.1 Test and Evaluation Master Plan (TEMP)

The Test and Evaluation Master Plan (TEMP) is the top-level T&E document used to support the Acquisition Process. Approval is required by Milestone I in programs using the legacy Milestone system and by Milestone B for those using the new system institutionalized in October 2000. The TEMP is the controlling document that describes and documents the major T&E events for an acquisition program. For all ACAT programs, it reflects and expands upon the program requirements defined in the ORD. ACAT I TEMPs are approved by OSD (jointly by DT&E and DOT&E). ACAT II TEMPs are approved by ASN(RDA), but can be delegated to the MDA. ACAT III TEMPs are approved by CNO N091. In NAVSEA, ACAT IV TEMPs are approved jointly by the SYSCOM Deputy Commander/PEO and COMOPTEVFOR; ACAT IVM TEMPs are approved by the SYSCOM Deputy Commander or PEO. This chapter provides an overview of the TEMP and offers some general guidelines for TEMP preparation, including the identification of common TEMP errors and deficiencies.

3.1.1 Purpose

The TEMP's purposes are to define and control the accomplishment of adequate T&E to support major program decisions, to identify special T&E resources and requirements to facilitate long-range planning, and to document the interrelationship between DT and OT events. The TEMP should be factual and specific, avoid generalities, and emphasize quantification. It includes quantifiable and testable requirements, both operational and technical. The TEMP describes the amount and type of testing to be conducted before each milestone, and the

resources required. The TEMP is a dynamic document to be reviewed periodically and revised when necessary.

3.1.2 Responsibilities for Preparation

The program manager is responsible for the TEMP and the preparation of all its parts, except Part IV and Part V, which deal with OT&E resources. In developing the TEMP, the program manager should establish early liaison with OPTEVFOR to ensure an integrated approach to TEMP planning through Integrated Process Teams (IPT). The IPT develops the TEMP and brings it to a level where appropriate approvals can occur without further formal review. Issues are resolved early, and consensus is reached in a timely manner, prior to TEMP signature. The IPT members are action officers from the major organizations involved in planning, coordinating, approving, and implementing the T&E program (suggested "notional" membership is discussed below).

For program using the Milestone system institutionalized in October 2000, since a TEMP is not required until Milestone B, an "evaluation strategy" is to be prepared soon after Milestone A to describe how the capabilities in the Mission Need Statement will be evaluated once the system is developed. Such a strategy should show how DT&E and OT&E will be integrated and what T&E will support the LRIP and full production decisions, in accordance with that program's overall acquisition strategy. For programs under OSD T&E oversight, this strategy is to be approved by DOT&E and the cog OIPT leader within 180 days after Milestone A approval. The program manager should prepare it in coordination with OPTEVFOR.

3.1.3 The TEMP IPT

If a T&E IPT has been established to develop the TEMP, it should include representatives from the Program Office, NAVSEA T&E Office, OPNAV T&E Office (N091), COMOPTEVFOR, Lead Technical Field Activity, and DT&E/DOT&E representatives if the program is under OSD oversight. Other participants, including the OPNAV Sponsor, Fleet Representatives, Range managers, Target sponsors, Support codes, and the Development Contractor should augment the IPT as necessary when related issues require their participation. Early and regular communications among members during program planning is critical in facilitating a mutual understanding of T&E objectives and limitations, and in identifying test resource requirements. The T&E Manager, as leader of the IPT, must ensure that continuous and in-depth two-way communication exists between

them and other Program IPTs to ensure a “systems engineered” approach to the development.

A T&E strategy should be prepared, showing the integration of DT&E and OT&E, consistent with the acquisition strategy.

The T&E Manager, in consultation with program, field activity, and T&E Office staff, prepares an initial draft TEMP to facilitate discussions. A final version of the TEMP should have the concurrence of the IPT before approval. Figure 3-2 outlines the process for developing the TEMP within the IPT structure, and for submittal for approval.

3.1.4 Timely Submittal and Updates

Timely approval of the TEMP and TEMP revisions are important to the success of the program. The requirements for Fleet ship and aircraft services—to support T&E events must be included in an approved TEMP in order for them to be addressed in the applicable quarterly Fleet scheduling conferences. TEMP approval is required to support Milestone II in legacy programs, and Milestone B in newer programs structured in accordance with the DoDI 5000.2 issued in October, 2000. Thereafter, a program manager should keep the TEMP current, reflecting the overall T&E strategy. Any major change to a program that requires a revised/updated Acquisition Program Baseline (APB) or Operational Requirements Document may also necessitate a TEMP revision.

3.2 Format and Watch Areas

The TEMP format is prescribed in DoD 5000.2-R. While the overall format has been standardized, the TEMP must be tailored to suit a particular system. The PM should do what makes sense. The format should be adhered to as much as possible because many organizations use the TEMP to extract selected information. The cover page contains lines for review and approval signatures. Part I includes mission description, system threat assessment, Measures of Effectiveness, and Suitability, system description, and critical technical parameters. Part II is entitled Integrated Test Program Summary, and includes a program management description and an integrated schedule chart that provides an overview of the major acquisition and T&E events (See Figure 3-1), as well as the funding profile. Parts III (Developmental Test and Evaluation Outline) and IV (Operational Test and Evaluation Outline) describe, in quantitative terms, the

scope of each major test period. Part V, the T&E Resource Summary, identifies the special resources required for the test program, when those resources are needed, and how much funding is required. DoD 5000.2-R has specific criteria on each section. Below are items within each section for which the PM should look out for based on this office’s experience.

3.2.1 Cover Page

The cover page varies slightly depending on ACAT of the program. Although the standard format implies a specific order for approval of the TEMP, the reality is that COMOPTEVFOR is the first to sign after submission by the PM. This allows minor changes to be incorporated after their signature, if needed, but before other major signatories. The signature cycle progresses from Program Manager to COMOPTEVFOR, then back to the DEPCOM/PEO. The TEMP is then forwarded to CNO (N912) for further processing. Additional details are provided in SECNAVINST 5000.2.

3.2.2 Part I

Part I is the System Introduction containing paragraphs A through E.

Paragraph B, System Threat Assessment - This section should focus on what aspect of the threat the system is intended to address. Most Threat Assessments are broad in nature and address

The more time that is invested in using the TEMP and other key documents to resolve questions about system performance requirements in advance, the less time will be required later to address issues in interpreting T&E results.

multiple areas. As such, your discussion should focus the reader on what specific, threat, levels, periodicity, and tactics are applicable.

Paragraph C, Measures of Effectiveness and Suitability (MOE/MOS), must contain clear caveats outlining the conditions by which these measures will be valid in a test sense. These caveats are reflected as notes to the MOE/MOS. Examples of such notes are:

- The environmental conditions in which the measures were derived and statements that any changes would require recalculation using agreed-upon criteria

- A listing of subsystems or functions covered by a particular measure (i.e., the level of Interoperability)
- Definitions of items such as "Available," "Failure," "Fault," "Degraded Performance," and Uptime"
- A description of, or reference to, the mission profile that the measures reflect.

These sections will require a significant amount of attention, but it will be well worth the effort. Getting this right the first time will avoid future conflicts when interpreting what data is valid for assessing the MOEs and MOSSs.

Paragraph D, the System Description - This section should clearly define and create boundaries around the program, including the overall level of Interoperability. It is important to outline which functions and equipment are part of the system itself, and to describe the key interfacing systems or functions that must exist for your system to meet its requirements. OPTEVFOR is tasked to assess the entire system with an "end-to-end" methodology.

Without a clear description of which interfacing systems are covered, conflicts will arise.

Paragraph E, Critical Technical Parameters

- This section lists the specific key hardware and software measures and their thresholds in a matrix format. These parameters and thresholds should also contain caveats describing under which conditions these measures will be valid in a test sense. These caveats are reflected as notes to the CTPs. Examples of notes are:

- The environmental conditions, such as sea-state or atmospherics, in which the measures were derived, and that any changes would require recalculation using agreed upon-criteria
- A listing of subsystems or functions covered by a particular measure,
- Definitions of such items as "Available," "Failure," "Fault," "Degraded Performance," and "Uptime."

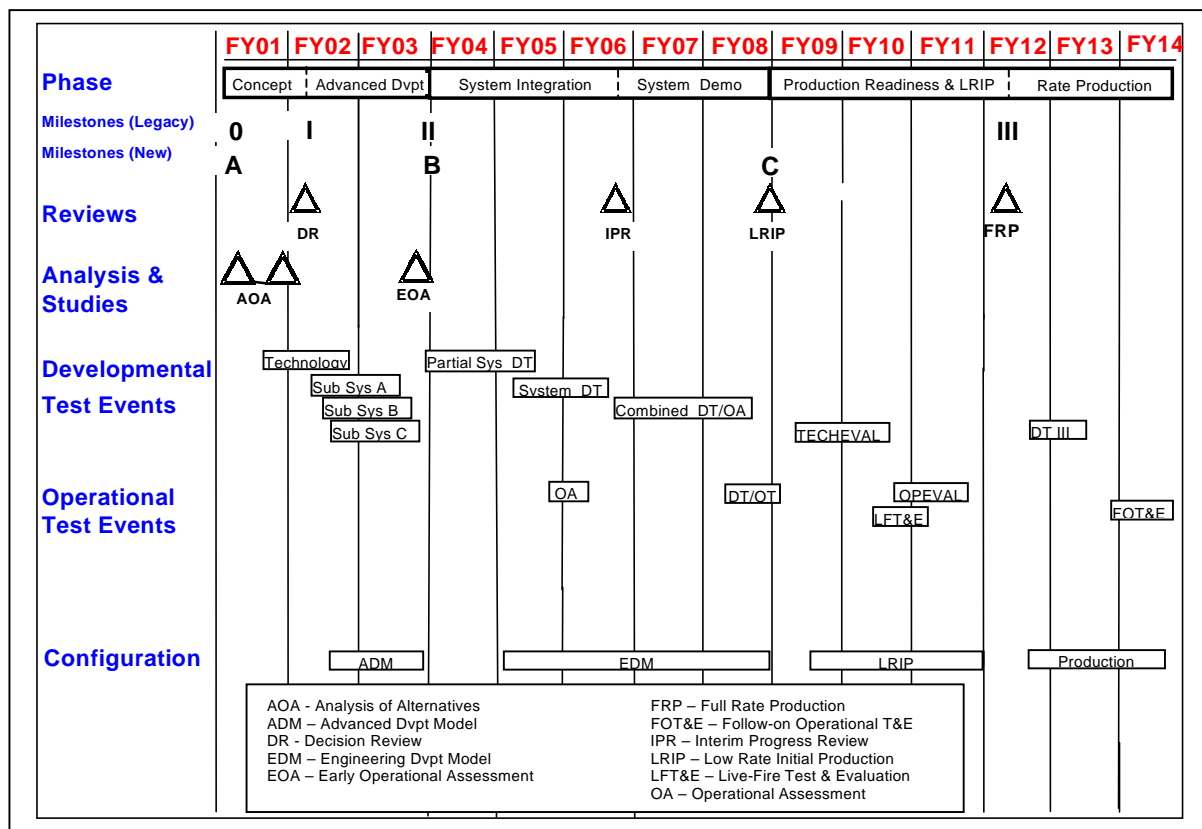


Figure 3-1 Notional TEMP Part II Integrated Schedule

Technical (or DT&E) and operational (OT&E) thresholds are specific performance requirements against which the system will be evaluated. Thresholds are the minimum acceptable values, consistent with operational requirements, which must be demonstrated for approval of program continuation. To the extent that these requirements are loosely defined, OPTEVFOR's evaluation will be necessarily subjective. The program manager and the Operational Test Director (OTD) could have different views as to the system's function. This has been the cause of several OPEVAL failures in the past. The program manager and OTD must develop quantifiable, testable parameters for inclusion in the TEMP. CNO is then responsible for providing thresholds.

Parameters and thresholds should be limited to those fundamental to the essence of system's requirements. For example, consider a missile system: its MOP thresholds should cover such items as Probability of Kill. The missile's MOP would cover such items as maximum range, kill radius, and maximum altitude. Characteristics such as fin turning rate, thrust ratio, or detonation time are better left to the CTP or the detailed system specification. There will be some overlap between DT&E (CTP) and OT&E (MOE/MOS)

thresholds. In general, an OT&E threshold reflects a mission outcome and the DT&E threshold the critical elements necessary to meet the OT&E parameter. A more detailed discussion on parameters and thresholds is presented in Chapter 2.

If necessary, an Appendix or Annex may be created to clearly outline the methodologies and measurement strategies used to calculate success or failures.

3.2.3 Part II

The Part II Integrated Test Program Summary, Figure 3-1, is an example. Most of the errors that appear in this chart in various TEMPs are the result of missing information, or errors in sequencing, such as tests beginning before test articles are available, test reports due before completion of tests, or a milestone decision prior to the scheduled completion of the tests which provide the results to support the decision. Undue optimism frequently appears with a milestone decision shown concurrent with the completion of testing. The schedule should reflect a realistic

schedule, allowing for document preparation and the necessary review cycles. The TECHEVAL and OPEVAL schedules should be clearly identified with some time in between for assessing data. Footnotes can and should be used, particularly to clarify anything that may be out of the ordinary.

3.2.4 Part III

Part III is the DT&E Outline. This section should clearly define which configuration or functions will be available to support a specific test event or phase. This section should also contain discussions on the processes by which surrogates, simulations, and models will be used and validated. Additionally, there should be clear descriptions of which threats surrogates will be used. If it is unknown at that time, then reference to a threat surrogate working group or similar forum may be required.

3.2.5 Part IV

Part IV is the OT&E Outline and is provided by OPTEVFOR. During Phases I and II, a demonstration of some test requirements might be met by concurrent DT&E and OT&E, or the use of DT&E results in OT&E, to reduce total test time and cost. This must be agreed to by OPTEVFOR in advance and documented in this part of the TEMP. Close coordination with the OTD will be necessary to ensure adequate integration of OT&E with the rest of the test program, and into the rest of the TEMP. In any case, a thorough review of Part IV is necessary to determine if their plan is executable within program schedule and fiscal constraints, and if the Critical Operational Issues (COI) can be resolved with the system as it has been described. These COIs are a series of questions outlined in PART IV, and form a bridge between the absolute nature of thresholds and the subjective assessment of operational relevancy. The term “resolving COIs” is used by OPTEVFOR to link all objective and subjective elements of a test program into conclusions of Operational “Effectiveness” and “Suitability.”

3.2.6 Part V

Part V is the T&E Resource Summary. The entries in this summary should reflect, where applicable, threat surrogates, type of aircraft, number of flight hours required, and the class ship, or, if known, the specific ship, the number of ship days, the type service required (dedicated, concurrent, or not—to-interfere basis), and the range resources required. The T&E Resource Summary should also include installation and removal schedules for special equipment. If OPTEVFOR or Fleet personnel travel to receive training, there are two required entries. The personnel training entry should indicate individuals, by rank/rate/grade, destination, and duration of travel. The planned travel entry reflects travel costs (direct travel and per diem) that OPTEVFOR and the Fleet must program and budget. The entry should be in dollars per fiscal

year. Common errors in this section tend to be omissions and inconsistencies in the sites and test surrogates chosen.

3.3 Abbreviated Acquisition Programs (AAP)

Special allowances have been established for Abbreviated Acquisition Programs (AAPs). While a formal TEMP is not mandatory, creating a TEMP will support acquiring Fleet Services for the test program. Alternatively, a PM would do well to outline the overall test strategy, measure of success, and resources required to do the job as in a TEMP-like document, such as a master test plan. This master test plan provides the skeleton for the program structure for hardware and or software, which is either commercially available or is based on systems in use by other services or foreign navies. These AAP programs are generally on the fast track, with limited resources available for extensive documentation. Nevertheless, there is a need for a top-level Master Test Planning document that outlines the objectives and the responsibilities of each major player.

3.4 Review, Processing and Approval

Developing a TEMP is the responsibility of the Program Manager and OPTEVFOR, through the test manager and in concert with the T&E Integrated Process Team (IPT). These IPTs are a good forum for developing and implementing a solid TEMP, chaired by the Program T&E Manager. To focus the initial T&E IPT, the test manager would do well to create a strawman TEMP, based on program constraints, with a proposed strategy. Deviations from this initial plan would be done in a collaborative fashion with all interested parties. The final version of the TEMP should be concurred with by the IPT before being submitted for approval.

A final version of the TEMP should be the outcome of a successful T&E team approach, and should have the concurrence of its participants before forwarding for final approval. For all TEMPs except for ACAT IVM programs, the approval cycle begins with a signature by the PM, who then forwards it to OPTEVFOR under formal correspondence. Once returned from COMOPTEVFOR, a formal signature package is created and sent to the cognizant NAVSEA Deputy Commander (DEPCOM) or PEO for signature. Upon signature, ACAT I, II & III TEMPs are forwarded to CNO N912 for the remaining approvals.

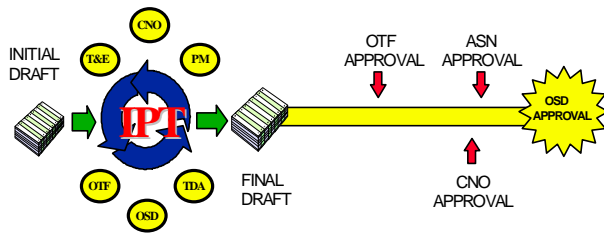


Figure 3-2
TEMP Development and Approval
Process

TEMPs for ACAT IVT programs are prepared in accordance with the same procedures in Paragraph 1, except: (a) they are not forwarded to CNO for formal review and signature; and (b) their approval is constituted by DEPCOM/PEO and COMOPTEVFOR signatures (i.e., CNO approval is not required).

TEMPs for ACAT IVM programs are prepared in accordance with the same procedures as Paragraph 1, except: (a) there are no OT&E thresholds in Part I and there is no Part IV, OT&E Outline, since COMOPTEVFOR does not conduct OT&E on ACAT IVM systems; (b) they are not forwarded to CNO or OPTEVFOR for review; (c) approval is constituted by DEPCOM/PEO signature alone; and (d) they need not be distributed to COMOPTEVFOR unless fleet services are required to support the T&E program. Following resolution, the program manager provides the TEMP, with concurrence by the NAVSEA T&E Office for the cognizant DEPCOM, whose signature constitutes approval.

After approval, all TEMPs are distributed to TEMP Annex A POCs and the NAVSEA T&E Office (SEA 91T).

Changes to the TEMP are usually made to address major schedule and/or threshold changes within a defined program. Revisions to the TEMP are conducted to address new phases of DT or OT supporting an evolutionary upgrade approach. TEMP Change Pages are published by the PM, with a signature page similar to that for a new TEMP. TEMP Revisions are treated as a new TEMP with a new cover sheet.

3.5 Other T&E Plans

As described above, the T&E Master Plan is more of a high-level planning guide and a contract between the PM and other interested test organizations than a vehicle for detailed test execution. Other types of T&E plans that program managers have found necessary and useful include:

- A T&E management plan for the overall program, to coordinate the activities of the many participants during the entire T&E program. Such a plan is discussed in Chapter 5.
- A T&E plan for a particular program phase. Coordinates the tests to be conducted at specific sites during a unique test phase. Chapter 5 discusses the "Integrated T&E Plan" (ITEP).
- Detailed plans for major test events.

Thorough, stand-alone, unambiguously written reports are an absolute necessity! They are a vital communication medium within each organization and among the organizations involved in the program.

3.6 Test Reporting

At every level of the organization, from the engineer in the lab and the technician on the ship, to the manager in the program office, many incentives minimize formal written reporting. Formal written reporting is expensive, time-consuming, and its full value and use are rarely understood by the developer of the report. However, formal reporting is an absolute necessity! It is a vital communication medium within each organization and among the organizations involved in the program. Its publication is a statement of the technical integrity of the events and analyses it describes. It is also a reference point for future corrective action. The program manager must cite test program results at many reviews, such as Milestone Decision meetings, Interim Progress Reviews (IPR), and the OPEVAL readiness review. The PM must provide some test reports to OPNAV and must even report some results to Congress in annual Congressional Data Sheets. Detailed test reports must be available to substantiate the conclusions and recommendations the program manager makes in these reviews.

Test reports are also important historical documents. Frequently, test reports from events conducted many years earlier are recalled during system production or during the development of engineering changes to the system. The problem for many smaller ACAT programs is that properly documenting test results is time- and resource-intensive. The immediate

need may not be clear, but the long-term benefits in less retest more than compensate for the resources necessary. In many cases the DT reports, including TECHEVAL report, are the only historical references for the record of what was conducted in the program. It is important for successive management personnel to understand the decision and tradeoffs made during development. The characterization of system performance is made not only from the testing being done today, but also from what was done in the past.

There are two types of DT reports supporting the development process; DT Event Reports and DT Summary Reports.

DT Event Reports - These are specific test reports written to support the engineering process. They address the “who,” the “what,” the “when,” and the “where” of a particular test event. Event Reports conclude with an assessment of the problem areas and risks.

TECHEVAL Reports - The TECHEVAL report is a DT Summary Report of the final System Development & Demonstration tests before the Independent IOT&E event (OPEVAL). It should list not only the types of tests conducted and their results, but also how closely the tests covered the threat and simulated real operational tests, and how the installation, manning, and training simulated what will be done in OPEVAL. The TECHEVAL report, supported by the DT event reports, should clearly show how all TEMP thresholds have been met, and that the program is ready for OPEVAL.

DT Summary Reports - These are written, in some cases for large, multi-phased T&E efforts to support the transition between Milestone phases or major program review points. The DT Summary Report summarizes all testing addressed in individual DT Reports to provide a sense of the scope and adequacy of DT. The Summary Report also relates the results of these individual tests to the specific Milestone exit criteria and/or TEMP thresholds, and determines the risks for the next test phase.

The numbers, types, and schedules of written T&E reports need to be planned well in advance. The program manager must receive T&E reports that will provide the proper information for making the major cost, schedule, and technical decisions during the program, to provide direction to the T&E efforts. The program manager must, in turn, provide summary results of the T&E efforts to higher authority at status briefings, milestone decision points, and budget reviews. As an example, in the case of selected ACAT I programs, the program manager must provide a brief summary of T&E results during the preparation of the Selected Acquisition Report (SAR) that becomes a part of the President's biannual budget submission to Congress.

Test results should not be released prematurely based on partial evaluation. This will always cause confusion and, more often than not, embarrassment. Letting Headquarters wait a few days until the results have been determined and confirmed is better than apologizing for incorrect or misleading reports later.

3.7 Updating Program Documents

There is a critical need to keep the system performance requirements and evaluation criteria that appear in the program documents (ORD and TEMP) current. The nature of the development process dictates that original requirements – and the subsequent documentation – be reassessed. The program manager should conduct at least an annual review of the TEMP to ensure that it is current. During R&D, engineering tradeoffs, which must be made frequently, will have impact on program documents. Such changes can usually be readily approved; however, these approvals must be reflected in the ORD and the TEMP, as it is against the requirements in these documents that OPTEVFOR evaluate the system. Also, any necessary changes should be approved as soon as the need becomes apparent. Immediately before TECHEVAL or OPEVAL is not the time to be requesting re-evaluation of system performance thresholds. The reviewing officials will rarely approve such changes at that time because it might appear that the changes were being allowed only to keep the program on schedule and within cost, in violation of the milestone procurement policies.

The ORD is one document that must be re-validated prior to each subsequent milestone decision. As such, great care must be taken to ensure that the ORD parameters and thresholds are still valid, based on the development and test results. Timing is also a consideration, since the results of DT may be the reason for updating an ORD. The revalidation process can average between six and 14 months, and will impact the scheduling of Milestone decision meetings.

The systems engineering and the T&E results may warrant changes to performance parameters and thresholds in the ORD. If so, those changes should be immediately.

The following are examples of problems that out-of-date documents create:

One program, begun as a relatively simple reliability improvement for an in-service system, ended as a completely new development that finally went to OPEVAL almost 25 years later. Early tests had been conducted with limited or no instrumentation. These same early test results were simply extracted from the test report and later proposed for inclusion in the TEMP as required operational performance characteristics. However, under scrutiny, they did not match recent performance predictions made with the latest knowledge in the field, and certain characteristics were not repeatable in the more modern versions of the system. OPEVAL failure would have been certain, had the program manager not carefully reviewed the proposed TEMP write-up.

In another program, it was discovered during system TECHEVAL that the performance fell substantially short of the TEMP threshold in one parameter related to the physical endurance of the personnel. Tracing the origin of the threshold value, the program manager found that it was proposed several years ago by medical personnel who thought the TEMP a convenient vehicle to promote "pushing the state-of-the-art" in this area. The threshold had not been questioned when the TEMP was staffed. As a result, the program manager found himself in the uncomfortable position of trying to get a performance requirement changed immediately prior to OPEVAL. OPNAV did approve the change, but this extra step delayed certification of readiness for the OPEVAL. It also jeopardized the OPEVAL itself, which had a limited schedule window due to the availability of Fleet support.

A third NAVSEA acquisition program, begun prior to the issuance of the current T&E policies and procedures, had as its objective a repeat buy of an in-service craft with a different, larger engine. No new R&M requirements had been planned or budgeted. When the TEMP was prepared late in the program in order to plan for an OPEVAL directed by the CNO, there was pressure by some OPNAV and subordinate command staffers to identify R&M performance thresholds (the pressure from subordinate command staffers was a reflection of the

recent heavy emphasis on R&M by their commander, and the pressure by the OPNAV staffers was to ensure that OPTEVFOR had adequate direction for assessing the system's operational suitability). The staffers even proposed using the R&M characteristics from the in-service systems as thresholds since these apparently would not pose any risk to the OPEVAL. NAVSEA successfully resisted this pressure because: (1) any statement of R&M thresholds would mistakenly imply that R&M had been an engineering consideration in this program; and (2) the R&M characteristics of the in-service systems were not definitely known. To provide the necessary guidelines for OPTEVFOR, OPNAV did publish "measures of effectiveness" separately from the TEMP for OPTEVFOR to use in their evaluation of the system's R&M, but these purposely were not identified as "thresholds" against which program success would be evaluated.

The lesson in all these examples is that requirements documents must always be kept updated throughout the program development. The TEMP is not the place to correct a fundamental requirements deficiency.

CHAPTER FOUR EARLY T&E

4.0 Introduction

This chapter covers the T&E conducted before the system enters full-scale development and test. This period extends from Milestone 0 through Milestone II of the Legacy process and Milestone A through Milestone B of the New process. Under the new process, there is no “program” identified. This “Concept and Technology” period is less formalized from a T&E standpoint, but it plays a critical role in that it serves as the principle building blocks of an ACAT I or II program, involving substantially new technology or a high-value, large, platform or platform-level, complex combat systems integration effort. The concept that is to be matured and sent into System Development depends on efforts such as Advanced Technology Demonstrations (ATD) and Advanced Concept Technology Demonstrations (ACTD) to examine concepts and identify potential risk or application areas and which alternatives and specific concepts are to be pursued.

Concepts that are generated and selected for continued evaluation are verified using some early hardware and software models. Program risks and uncertainties are identified, and some are resolved, and the system's allocated baseline configuration and other documentation necessary to enter System Development are prepared. Most of the operational and technical measures are identified during this period. These demonstrations become more formalized with the development and assessment of subsystems before being integrated into a system. The objectives of dedicating a period specifically for technology maturation are to formulate feasible program concepts and examine the alternatives; to identify risks and plan appropriate risk management; to identify preferred technological alternatives and to develop plans for conducting sub-system development.

4.1 Defining the Concept

Development of possible concepts to fill an operational need is generally one of multiple iterations, involving a series of engineering investigations and trade-off analyses reflected in Analysis of Alternatives (AOA) that consider required system functions, available technologies, Integrated Logistics Support (ILS) considerations, and a first cut at life-cycle costs for each identified concept. Little, if any, testing is usually conducted during CE and generally entails “systems analyses,” which include Modeling and Simulations (M&S) to predict what could be achieved by the system.

The desire for new capabilities can evolve in response to a new threat, a new operational need, or the opportunity to capitalize on new or existing commercial technology that has become available. If use of an existing military or commercial system, or a minor modification of an existing system, cannot provide the new capability, a new acquisition program (either conventional or NDI) is initiated and progresses to Milestone II (Legacy) or B (New). An Analysis of Alternatives (AOA) is generated to support this transition in response to the Mission Need Statement (MNS) generated from the results of the concept studies. The alternative to be pursued is selected, and an Initial Operational Requirement Document (ORD) is issued. The program is included in the Program Objectives Memorandum (POM) for budget purposes. Operational and technical parameters are defined, but often they are not assigned quantitative values until after Milestone II/B. These parameters and estimated performance measures are usually the subject of trade-off studies that attempt to optimize performance within cost and schedule constraints.

4.2 Early Testing

When Navy does testing early in the conceptual period, it is not likely to be on a prototype of the system, but rather, on an experimental component that may be the heart of the development effort. Testing is limited in order to devote no more resources than necessary to the process of selecting a concept. This testing also may consist of adapting a current Fleet system to investigate its potential for another application. If there is any involvement by OPTEVFOR, it is only to conduct an early operational assessment on prototypes to assess major risk areas in support of the transition to System Development. If test reports are developed, they are usually done by a Navy Technical Development Agent (TDA). Generally, a TEMP is not necessary at this point, nor is it required until after Milestone II/B. However, for larger programs under OSD oversight, a defined T&E strategy is required.

DoD uses many different methods to advance from the prototype realm to something that reflects what can be produced just after the production decision. For larger ACAT programs, this can be a very formalized T&E program of prototypes with specific exit criteria at Milestone II/B. For other efforts, it can be early field activity demonstrations of applied technology, or new uses of existing or older technology. Many efforts today are an amalgam of all of these approaches.

One example of a unique approach is the Advanced Concept Technology Demonstration (ACTD) used to demonstrate application of new or unique uses of existing technology to meet a military requirement. The ACTD demonstrates a viable concept that can be introduced into the Fleet on a very limited basis for a two-year period. Should the concept be proven, a formal acquisition program may be initiated somewhere in System Development. Though no formal T&E (TEMP) program is initiated during ACTD events, an assessment program is established to objectively evaluate ACTD adequacy. No matter which vehicle used to demonstrate technology maturity, it is the results of T&E that influences the transition into System Development.

4.3 OPTEVFOR's Involvement.

Although OPTEVFOR involvement during the early phase is limited, the program manager should establish contact with the Operational Test Director (OTD) during the latter part of the Sub-system development before Milestone II/B. The OTD can assist in defining operational performance measures, in providing valuable operational insights into the concepts being considered, and in projecting how OPTEVFOR would test them in later phases of the program. In particular, the program manager and the OTD should determine whether special targets, range instrumentation requirements, or models and simulations need to be planned and budgeted at this time. The closer the program manager works with the OTD throughout the program, the better the chances of having a more effective, efficient, and cohesive T&E program.

4.4 Early Operational Assessments

The OSD Director of Operational Test and Evaluation (DOT&E) has institutionalized EOAs to provide inputs from the Services' Operational Test Agencies to early milestone decisions and design reviews, when actual OT&E cannot be conducted because of lack of hardware in a suitable configuration. An EOA is an evaluation by OPTEVFOR on the potential of the program to achieve the system's operational performance requirements. They are distinguished from operational test reports by the fact that they are not based on actual operational tests. Instead, they are based on available results from development testing, or simulation and modeling results, or both. If COMOPTEVFOR feels that the available information is not adequate to support a projection of whether or not the system will meet its operational requirements, he will so report. And, instead of a projection, he will provide his observations of test results and recommendations for system improvements. EOAs are provided to support Progress Reviews

and the Milestone decision. An EOA is usually required only on ACAT I and other OSD-monitored T&E programs. An example of one was the EOA for the Virginia Class submarine, where COMOPTEVFOR projected the operational effectiveness of the entire submarine system, including its Combat System, based on Models and Simulations

4.5 Early Development Examples.

A good example of T&E during this period is provided by the Remote Minehunting System (RMS) program. Navy and industry collaborated to produce a prototype undersea autonomous vehicle to detect and classify mines. The vehicle was created by merging new and existing technology to form a test vehicle that examined the autonomous mine-hunting missions and which functions should be part of the future requirement. The results of at-sea testing proved the concept to have merit to support the creation of an ORD, and that the technologies were sufficiently mature to allow entering System Development.

After more than five years in exploratory development and technology assessment efforts, the Mk 50 Torpedo evolved with two approaches and two contractors. The demonstration took four years, and consisted of contractor design, fabrication, demonstration, and evaluation of basic components, subsystems, systems, and full-scale prototype torpedoes. A little over halfway through the demonstration, NAVSEA decided to continue with only one of the contracts because it involved a substantially lower risk design than the other.

Smaller programs will tend to have less formal technology assessment events where limited Navy testing does take place. Most of this time will be spent integrating existing systems to demonstrate a warfighting capability. Sometimes the objective is not to select a concept, but more simply to demonstrate the feasibility of a known concept or technological alternative. This usually takes the form of "black box" testing at a Navy laboratory, when a single performance attribute is in question. Some examples of programs where testing of this type has taken place are the Electrically Suspended Gyro Navigator, the Shipboard Data Multiplexing System, and the Doppler Log. In the Doppler Log program, testing was used to further allow selection of the operating frequency of the system, thereby bounding some of the design parameters to be addressed in further development.

The scope of effort is driven primarily by the extent of new engineering development associated with the design. Generally, programs

that use current technology have technology maturation phases and may even be combined with sub-system development. In others, the design, fabrication, and testing of one or more Advanced Development Models (ADMs) will be necessary.

The Advanced Combat Direction System Block 1 program conducted DT on prototype software. The evaluation was conducted at the prime contractor's facility. OPTEVFOR witnessed this testing and prepared an Early Operational Assessment (EOA) of the system to support the transition into system development.

The WLY-1 Sonar Intercept system conducted a series of DT tests using a number of different prototype arrays arranged along the hull of a submarine. These arrays were provided by contractors and navy field activities to assess performance requirements, a critical step in assessing how much processing capability was required for the WLY-1 inboard equipment. The arrays varied in frequency coverage, placement, size, and physical modularity. After 18 months of sub-system development and dockside and at-sea testing, the optimum balance between size, placement, frequency range, and number of arrays was found. The results allowed the overall WLY-1 design to proceed to system development with significantly fewer array footprint and processing requirements than had been envisioned.

During early design in the LPD 17 Amphibious Assault Ship Program, several EOAs were conducted to support the equivalent of a Milestone decision to commence with lead ship detailed design and construction. The EOA consisted of review of preliminary design and planning documents. Areas assessed in the EOA included the following: amphibious operations, aircraft operations, cargo handling, combat systems, C4I, survivability, interoperability, manning, human factors, and others. COMOPTEVFOR recommended numerous design changes to the ship, and they were considered in the detail design. DT included use of design models and mock-ups to verify design requirements.

In the Strategic Sealift Ship Program, two EOAs were conducted. These EOAs consisted of one assessment of conversion ship designs; the second EOA was an assessment of the new construction ship designs. Of note, this program's OT&E was a multi-service OT&E (MOT&E), with Navy as lead OT test agency teamed with Army OT testers. To evaluate cargo-handling capabilities of the various ship designs, a Roll-On/Roll-Off (RO/RO) rate model was developed for the PM as a design tool, and was evaluated by OPTEVFOR.

4.6 Test and Evaluation Identification Number (TEIN).

Shortly after formal initiation of the program and upon approval of the ORD, the program manager should obtain from the CNO the assignment of a Test and Evaluation Identification Number (TEIN). The program manager should initiate this request by submitting a letter with the information shown in Figure 4-1. OPNAV responds by promulgating a TEIN assignment letter to the program manager and OPTEVFOR, assigning the TEIN, formally establishing the T&E program.

[SSIC]
[orig. code]
[date]

From: Commander, Naval Sea Systems Command/ or PEO
To: Chief of Naval Operations (CNO N-912C)
Via: Chief of Naval Operations (Program Sponsor)
Subj: REQUEST FOR TEST AND EVALUATION IDENTIFICATION NUMBER
Ref: (a) SECNAVINST 5000.2B
(b) Operational Requirements Document

1. In accordance with reference (a), it is requested that a Test and Evaluation Identification Number (TEIN) be assigned to the [Official Title] Program (Program Element -----; Project Number -----). [Project Description ----- (see notes 1 through 3)]. A Test and Evaluation Master Plan (TEMP) will be approved prior to requesting/Fleet services, including ACAT assignment.

2. Points of contact are:

a. Requirements Officer [i.e., Sponsor]: [NAME], [CODE], [AV NUMBER]

b. Developing Agency [i.e., Program Manager]: [NAME], [CODE], [AV NUMBER]

c. T&E Coordinator [i.e., N912 P.O.C.]: [NAME], [CODE], [AV NUMBER]

3. Milestone status: [Provide actual or projected dates of milestone decisions.]
M/S A M/S B M/S C FRP [if anticipated]

Copy to:
CNO [Requirements Officer, T&E Coordinator]
COMOPTEVFOR (10, OTC/OTD)

NOTES:

- (1) By endorsement, the Sponsor will confirm that the request for TEIN assignment is supported by a valid Operational Requirements Document (ORD); the ORD should be referenced in this description.
- (2) A one or two line project description along with the ultimate user(s) should be included to further identify the project.
- (3) Requests for TEIN's which are to be assigned as "F" or "K" projects, as defined in SECNAVINST 5000.2B, may delete required information as appropriate.

Figure 4-1. Format for TEIN Request

4.7 The TEMP's Role

The Test and Evaluation Master Plan (TEMP) is the fundamental top-level T&E strategy document required by DoD and Navy in support of Acquisition programs. A detailed discussion of the specifics of the TEMP is covered in Chapter 3; only its applicability to the Milestone decision process is covered here. All programs are required to have a TEMP at this point with defined test parameters and thresholds in place. For ACAT III and IV programs (which normally do not begin until just before Milestone II/B), the TEMP is required around the time that this Milestone would occur, the beginning of the first fiscal year containing program funding.

TEMPs are required to support Milestone B, or its equivalent in a program that does not have an identifiable Milestone B. In any case, it must be approved prior to requesting fleet services to support an at-sea T&E event (which is usually nine months prior to the beginning of the quarter in which that event is planned). Aside from the requirement to have a TEMP, it is desirable to have one as early as possible in order to have a commitment from OPTEVFOR and OPNAV on the adequacy of the planned scope of T&E and the interrelationship between the DT and OT events. Since the negotiations and approval process for TEMP's can be time-consuming, it is

strongly recommended that the program manager begin preparing the TEMP as early as possible. The TEMP requires a coordinated team of those with acquisition T&E experience. An full description of the need and mak-up of a TEMP IPT is covered in 3.1. Initial versions of the TEMP will lack specifics, and will require later changes (e.g., system thresholds often evolve as the program matures and the results of design tradeoffs often cause threshold changes). Thresholds should be firmly established in an updated and approved TEMP at Milestone B, or its equivalent, even if it is expected to quickly become outdated in some parts. A note on TEMP for ship acquisition programs: TEMP are required only if the CNO has required Operational T&E by OPTEVFOR. OT&E may not be required for other non-combatant type ship classes ships, such as tenders, repair ships, fleet oilers, and supply ships, unless these are designated ACAT I Programs. SECNAVINST 5000.2B documents this exemption (Chapter I of this handbook also addressed ship acquisition programs).

4.8 Assessing Requirements

Models and simulations are used to evolve the top-level requirements examined in the concept stage to what is needed to further characterize the approach selected and pursued in system development. The EOA in this phase depends almost exclusively upon M&S to continue supporting high-level assessments of potential overall weapons and combat systems effectiveness. The objective is to transition the M&S used in advanced development to system development, and then to production, using a progressive layered approach. This continuum saves efforts in the long run by having the same models, with appropriate verification and validation, support the entire acquisition. During system development, M&S evolves from MOE/MOP derivation to detailed requirements specification derivation and assessment. M&S is used early on to help flush out interoperability issues for T&E, and then again during TEMP development, to determine the constraints and limitations of the T&E approach

4.9 Solidifying Thresholds

A threshold is the minimum acceptable performance value that a system is to achieve and still be approved for production and fleet introduction. There are more detailed discussions on Parameters and Thresholds in Chapters 2 and 3. Thresholds must be established by the CNO, based upon the minimum capabilities required for the system/equipment to perform its intended mission. Thresholds must not be based upon engineering estimates of a preconceived design or upon a manufacturer's promise. Rather, they should be chosen with great care, as program success will be judged against these values.

Thresholds are of value only when actual performance can be measured against them – in other words, they must be testable.

There should be as few thresholds as possible, but enough to determine achievement of each dominant system characteristic. Thresholds are of value only when actual performance can be measured against them - in other words, they must be testable. The most important of these thresholds are designated as Key Performance Parameters (KPP), as outlined in Chapter 2 and reflected in the Acquisition Program Baseline (APB). Their achievement must be linked to specific DT&E and OT&E events, during which system/equipment performance will be measured.

The establishment of separate thresholds for DT&E and OT&E is preferable. The program manager will determine the DT&E parameters and COMOPTEVFOR the OT&E parameters. The numerical values for DT&E parameters and OT&E parameters will be assigned by the CNO. The T&E working groups examine the ORD thresholds for testability and develop the threshold "notes" (caveats) that describe the test conditions or methodologies by which the ORD MOE/MOS are valid. This forum also outlines which surrogate target and other test articles will be appropriate to demonstrate that the top-level ORD thresholds have been met. Many factors must be considered before thresholds and caveats can be agreed upon. Items such as test site location, time of year, environmental conditions, adequacy of the surrogate, and quality of the data accuracy collection requirements become critical to assessing the affordability and executability of the T&E program. When necessary, the working group will recommend clarifications to the ORD descriptions for the sponsor to publish as a result of deliberations.

By Milestone B (II), every threshold value should be firm and if values cannot be established, the program manager should consider delaying the Milestone decision. Further, if established values cannot be fully supported within planned funding, the program manager should request additional funding or propose threshold modifications that would accommodate

the projected funding levels. At any point in time where the system performance requirements cannot be achieved within the funding levels, the program is "NOT EXECUTABLE", and must be restructured. In addition, changes recommended to TEMP Part I parameters and thresholds after Milestone B should be made only after first assessing the ORD.

4.10 Threshold Pitfalls

The following paragraphs outline some of the pitfalls and lessons-learned with regards to the selection, derivation and management of thresholds.

- ❑ Frequently, OPTEVFOR or OSD TEMP reviewers request that specific additional parameters or thresholds be placed in the TEMP. Although sometimes easy to accommodate, this type of request should be considered only after careful deliberations. The requested thresholds are sometimes very detailed technical characteristics that are more related to the design specifications than those needed to make acquisition decisions. It has been found that including verbiage in the DT&E section of the TEMP that makes a commitment that these characteristics will be tested, rather than including specific thresholds in the TEMP has satisfied this type of request.
- ❑ A "double jeopardy" can be avoided if thresholds are checked to ensure that they do not interleave effectiveness and suitability factors within their definition. Measures of Effectiveness are separate and distinct from Measures of Suitability. This problem occurs from time to time in weapons programs. For example, the system's overall effectiveness is stated to be equal to its ability to neutralize the target, multiplied by its launch reliability. This type of parameter description merges two subject areas that are deliberately separated. The result would be that a program would be penalized twice for the same fault.
- ❑ Change parameters and threshold definition in the TEMP as soon as it becomes apparent that a change is necessary. Here are some examples:
 - The AN/BSY-1 Submarine Combat System Mean—Time-to-Repair (MTTR) threshold did not take into account the time necessary to enter the sonar dome to repair transducers. To repair those items, it is necessary to evacuate the dome, which takes several hours and which, if added to actual repair time, would have caused the MTTR threshold to be breached. An agreement with OPTEVFOR was necessary to avoid miscalculation during OPEVAL. Ideally, this should have been included the TEMP.

- An operational threshold for Standard Missile magazine storage time before use was specified as 8 months in the requirements document. In order to test this, any missiles used before an operational test would have to be loaded on-board ship 8 months in advance. This proved to be impossible during development, due to ship scheduling difficulty and because of the need to make changes to the missiles as a result of DT testing, which would have entailed downloading the missiles and restarting the clock. This threshold was provable only after the missile was in fleet use and in an OT-III type of scenario. Hull mechanical and electrical (HM&E) systems typically have this type of threshold problem, where shelf lives on the order of years are typical. Since TEMP parameters and thresholds are to be "testable," parameters such as storage time should be avoided.

- Also in the Standard Missile program, a maximum altitude threshold was set based on the threat. However, demonstration of effectiveness of the proximity fuse in about the top 10 percent of that region could only be demonstrated by test missiles if there was direct hit. Range instrumentation did not have sufficient accuracy to confirm proper fuse action for a proximity contact unless there was an actual warhead in the missile. The TEMP threshold should have been set at the 90 percent% of altitude level, instead of the actual 100 percent.

- In the Damage Control Wirefree Communication System program, two new effectiveness parameters were requested after the design was complete. The OPNAV Sponsor agreed and supplied threshold values he felt the design could meet, but the system failed its OPEVAL because it failed to meet these two thresholds. Since the program was well past Milestone B equivalent (actually it had completed development), NAVSEA should have better resisted accepting the new parameters and thresholds.

- The surface ship vertical launch system (VLS) program manager wanted to raise the TEMP reliability threshold from 50 to 1500 hours mean—time—between-failures (MTBF) in his

TEMP revision. He felt that the 50 hours reflected poorly on the system and did not accurately reflect its true capability. Even though the test results supported the higher number, the lower threshold was retained because it was the minimum number the Fleet needed to support the total combat system operation. The latter is the true definition of a threshold and should be adhered to.

- For the AN/BQQ-5E Submarine Sonar system, RMA parameters failed to take into account that towed array handling equipment, as well as the array itself, was not to be repaired at sea. There was a major impact to assessing support maintainability and availability. It was found that since these subsystems could not be repaired without returning to an intermediate repair facility, the mission was essentially concluded with the first failure, with Ao unable to be calculated, maintainability was meaningless, and the Mean Logistics Delay Time was essentially variable. These conditions had a major impact on Operational Availability. Clearly, the intent of the ORD thresholds was to focus on those efforts, in which the submarine crew could reasonably be expected to control to maintain operational mission readiness while on station with supplies on-hand. Any mission-critical failure that could not be repaired at the "O" level would logically require the submarine to come off-station, thus ending the mission.

4.11 Exiting the Initial Phase

At the close of technology maturation and sub-system development, a Milestone review is held to address alternative design concepts, alternative acquisition strategies, expected operational capabilities, industrial base capacity, production capability, readiness, support, personnel requirement projections, and cost estimates. If use of an existing military or commercial system, or a minor modification of an existing commercial system, cannot provide the new capability, the results of studies and testing will indicate that an acquisition program (either conventional or NDI) will be initiated. An Analysis of Alternatives (AOA) is generated in response to the Mission Need Statement (MNS), generated. An initial Operational Requirement Document (ORD) is issued and the program is included in the Program Objectives Memorandum (POM) for budget purposes. A TEMP is then generated that includes an outline of the T&E program and performance thresholds to be demonstrated during System Development.

CHAPTER FIVE

T&E DURING SYSTEM DEVELOPMENT, TECHEVAL AND OPEVAL

5.0 Introduction

System Development represents the evolution from systems engineering to production readiness. From a T&E perspective, it is the characterization of inter- and intra-system performance necessary to demonstrate and document the system's technical adequacy, operational effectiveness, and operational suitability. The final output is a baseline configuration for production and fleet introduction.

There are two aspects of system development. The first is the integration of subsystems into a system that meets inherent performance requirements. The second is the demonstration of the system in an operational environment when interfaced with other systems. System Development reflects conceptual maturity, technical competency, and operational immaturity. There is an understanding of what is desired, the tools and resources to do the job, and some idea of how the product is to be used in the operational world. The end of development reflects the maturity of all three and T&E quantifies this through DT and OT. Many different risk reduction aspects of acquisition converge during development, in which T&E plays a major role. Most ACAT III and below programs will begin the acquisition process at this phase, based on earlier T&E of crude prototypes or of integrated sub-system elements of much larger systems.

Development testing is conducted at contractor facilities, Navy engineering activities, Land-based Test Sites, on-range, or open-ocean, to verify achievement of incremental performance requirements. These objectives reflect the sub-system level thresholds and expand to system, inter-, and intra-system requirements based on the sphere of influence of the system. T&E is performed to verify the achievement of improvements made as a result of system development, to verify the achievement of required technical performance specifications, and to identify any improvements for the production version.

Milestone C under the new process reflects the point where the system has sufficiently matured to enter Low-Rate Initial Production (LRIP). More formalized DT&E and IOT&E are conducted to support the decision for Full Rate Production (FRP). The final phase of DT&E prior to the full rate production (FRP) Decision Review is the Technical Evaluation (TECHEVAL), conducted in the system's intended installation environment under more operationally realistic constraints. For shipboard systems, TECHEVAL is usually conducted on an active Fleet ship at-sea. The final phase of OT&E prior to the FRP decision is the Operational Evaluation (OPEVAL),

conducted on a production-representative system in the intended operational environment. Details of TECHEVAL and OPEVAL are presented later on in this chapter.

5.1 The Engineering Development Model (EDM)

For system acquisition programs with moderate to high technical risks, system development often involves the fabrication and test of one or more Engineering Development Models (EDM), followed by the fabrication and test of one or more pre-production models. Programs of lesser risk normally have only EDM(s).

EDMs are prototypes of the complete system, and are functional equivalents of the system. They are built for one or more iterations of the build—test—fix—retest cycle, as such system attributes as reliability, maintainability, safety, and supportability are verified. An EDM can be considered to be production-representative, but may not have the exact physical configuration of the planned production system. For example, commercial or other non-militarized equipment is used in non-critical areas of the design, while other equipment and component designs are being validated. DT&E is performed on EDMs to reduce the design risks and uncertainties prior to fabrication of a more production-representative model (LRIP), and to verify attainment of technical performance objectives in the components, subsystems, interfaces, and finally, at the total system level. OT&E may be performed on EDMs to verify aspects of operational performance while design decisions are being made. Operational assessments (OA) are also routinely conducted on EDMs to identify risk areas during development and to support LRIP decisions. If dedicated OT&E is not conducted, OPTEVFOR will probably (and should be invited to) monitor DT&E to gain an early familiarity with the system's capabilities and limitations, and should be given an opportunity to comment on design tradeoffs. Additionally, in order to combine testing where appropriate, OPTEVFOR and the PM may want to be involved in DT to attempt to resolve as many OT requirements as possible.

Once a design is sufficiently validated, production-representative models of the system may be built to provide test articles that are as close to the final production configuration as possible. LRIP units are usually procured in limited quantities, specifically for support-required IOT&E testing leading to FRP. (A larger number of LRIP models may be authorized at Milestone B for a variety of reasons, such as urgent Fleet

deployment requirements, a producibility risk that must be resolved prior to Milestone C, or the need to minimize very uneconomical stand-down periods for production facilities and associated personnel for construction or test.) The number of EDM units manufactured during system development must always be justified against the increased financial risk of subsequent rework of the design and manufacturing flaws. The number of LRIP units is limited to 10 percent of the total buy, unless otherwise approved by the MDA.

5.2 System Integration

System Integration is used to reduce the integration risks and demonstrate sub-system maturity during the development. System Integration transitions into System Demonstration following an Interim Progress Review (IPR). The objective of both of these periods is to conduct sufficient testing to mitigate risk, and to support production of Engineering Development Models (EDM) and future Low Rate Initial Production (LRIP) units.

5.3 Testing During System Integration

The objectives of testing during System Integration effort are: (1) to verify that the areas of technical risk are resolved, (2) to assist in solidifying the technical approach during System Demonstration; (3) to verify that the system has, at least, the potential to be technically and operationally effective, as well as operationally suitable; and (4) to determine the T&E requirements to be achieved during TECHEVAL and OPEVAL. To accomplish this, DT&E may be conducted on each alternative system. If it is expected that the system will employ new operational concepts or if it involves significant operational risks, an Operational Assessment (OA) or even early Operational Testing (OT) should be conducted on the more promising systems; i.e., those that meet most of the performance objectives. If such Initial OT&E (IOT&E) is conducted, the NAVSEA Deputy Commander or PEO notifies OPTEVFOR and the CNO in consultation with the NAVSEA T&E Office, that the system is ready to support such T&E. This is a less formal review and certification process than is used for OPEVAL supporting FRP, which is described later.

DT&E and OT&E conducted on the candidate system(s) should consider performance, compatibility and interoperability, electromagnetic susceptibility, logistic supportability, and reliability and maintainability, to minimize risks. The vulnerability of the system to hostile weapon systems and ECM should be analyzed to the extent possible. Design features to enhance system survivability should be developed and tested. The preferred approach is selected from the comparison of T&E results and the cost/price analysis of further development, production, and service life. In some rare cases, more than one system can be chosen to continue through development when operational risks may still be high, but the alternative systems cannot be further proven without using at least a pre-production prototype model.

5.4 Role of the DT Report

A DT report, that summarizes the testing conducted to date and the technical risks is used to support the Interim Progress Review. As such, this report should address not only what was done and the results to date, but also outline where the risk areas are and the changes to the engineering process for progression into System Demonstration. The Report should also address special areas of study required to reduce the risks for LRIP.

5.5 The Interim Progress Review

The Interim Progress Review is the evaluation of the program's readiness to proceed from a series of integrated subsystems to an integrated system. Testing must demonstrate that the system is able to operate consistently with the ORD. The program manager should be prepared to answer these T&E related questions affirmatively:

- Have the technical questions and critical sub-system maturity issues posed at Milestone B been adequately resolved?
- Has the T&E conducted on sub-systems been sufficient to reasonably ensure that the performance requirements can be achieved?
- Have the technical questions and critical issues to be resolved prior to LRIP and or FRP been identified? Is the test program to resolve them adequately planned?
- Have the technical and operational performance requirements and thresholds been refined based on testing to date?
- Are the test program scope and schedule adequate for the System Demonstration phase?
- Are adequate funding and resources available to support all required T&E during System Development?

5.6 System Development

System Development is the most critical phase of the acquisition program. Because the results of SD provide information to support both the initial production investment decision (LRIP) and the Full Rate Production (FRP) decision, the T&E conducted is the most extensive and formalized. Test planning and reporting requirements are therefore the most rigorous of the entire program effort. The program manager becomes more deeply involved in T&E activities as the linchpin between the engineering/T&E team and the authorities reviewing and approving the program's entry into production.

A series of DT, combined perhaps with a series of Operational Assessments (OA) and Operational Tests (OT), is conducted to assess the performance and identify risk in increasingly operationally relevant scenarios. Program risks and uncertainties are identified, and some are resolved. System Development establishes the system's allocated baseline configuration and produces the documentation necessary to reach LRIP, and ultimately Full-Rate Production (FRP). The involvement of OPTEVFOR and other organizations in the T&E community, becomes more focused, and the documentation is more critical. The objective is to reduce program risk, ensure system supportability, design for producibility, and demonstrate system integration and utility.

The engineering development program of the surface ship Vertical Launching System (VLS) was very successful, in large part because of a comprehensive T&E approach that addressed key risks and uncertainties in a methodical manner, and because of the extensive use of land-based testing. The key document in this test program was the Integrated Test and Evaluation Plan (ITEP). The ITEP formalized the relationship between the VLS TEMP, the VLS Prime Item Development Specification (PIDS), the prime contractor's Master Test Plan, and eventually the Master Test Book. Specifically, the ITEP amplified the DT test events outlined in the TEMP, integrated Navy and contractor tests to maximize data yield, described how each key parameter would be evaluated in each event, and analyzed each specification requirement for proper system development.

- The test program manager selected the 10 key parameters critical to successful surface ship VLS performance in OPEVAL.
- Development testing was arranged so that:
 - Each parameter would be addressed in successively complex land-based tests prior to TECHEVAL.
 - Critical issues were addressed as early as possible.

- Test phases were scheduled to permit data gathering as soon as subsystems were available for test.

Eight contractor and eight Navy test phases were structured as building blocks. The initial tests were limited in scope and complexity, and the number of issues addressed increased in successive tests. The first 14 test phases were performed at land-based test sites, the contractor's plants, the Naval Weapons Handling Center, the White Sands Missile Range, and the AEGIS Computer Program Test Site. The remaining phases were conducted on a pilot production model installed in a dedicated test ship. Successful completion of many of the test events were prerequisites to others. A documentation tree was prepared, and demonstrated the relationships between the test plans and reports for each phase. Prior to each Navy DT test event, a Mission Readiness Panel, composed of high-ranking Navy officials and contractor executives, convened to review and approve proceeding into the tests. As a result of this thorough planning and rigid adherence to the plan, TECHEVAL confirmed previous test results and was a successful prelude to OPEVAL.

		TEST PHASE																		TEST PHASE (○ MEANS DATA GATHERED AS PLANNED AND RESULTS WHERE SATISFACTORY)			
		LAUNCH/CATCH	CON FLOW	DT/II A	DT/II B	STRIKE DOWN	ENVIRONMENTAL	DEMO	ACCEPTANCE	DT/II C	DT/II A	DT/II D	MAINTAINABILITY	DEMO	DT/II E	DT/II F	PPM/1	INTEGRATION	SYSTEMS			INTEGRATION	DT/II G
KEY PARAMETER	PARAMETER SOURCE																						
GAS MANAGEMENT SYSTEM PERFORMANCE	CRITICAL ISSUE	●	●	●	●											●					●	○	
LAUNCH ENVIRONMENT EFFECT ON MISSILE PERFORMANCE	CRITICAL ISSUE				●	●										●					●	○	
CANISTERED MISSILE PHST	CRITICAL ISSUE											●	●								●	●	○
STRIKEDOWN SYSTEM PERFORMANCE	CRITICAL ISSUE					●																●	○
ENVIRONMENTAL SUITABILITY	VLS SPEC							●												●	●		○
COMPUTER PROGRAM PERFORMANCE	VLS SPEC										●			●			●	●			●		○
INTEGRATION WITH MISSILE AND WCS	VLS SPEC										●			●			●	●	●		●	●	○
RELIABILITY, MAINTAINABILITY, AND AVAILABILITY	REQ TECH CHAR												●	●			●	●	●			●	○
REACTION TIME	REQ TECH CHAR												●				●	●			●		○
FIRING RATE	REQ TECH CHAR												●				●	●	●			●	○

KEY:	
●	= TEST PHASE ADDRESSING PARAMETER
●	= CRITICAL TEST PHASE

Figure 5-1. Test Planning Matrix (Example)

5.7 System Development Examples

The following examples demonstrate the myriad of T&E approaches used during System Development, and how prototypes and EDMs are used support the T&E process.

After an extensive phase of risk-reduction testing of a version of Standard Missile, 45 pre-production and 65 pilot production missiles were fabricated for engineering and manufacturing development testing. The test program used simulation, prototype radomes, warhead and fuse prototypes, breadboard missile receivers, missile shapes, and development rocket motors. Development testing was then conducted in six DT test phases and three OT test phases. Testing consisted of non-firing tests, and both land-based and at-sea firings of blast test vehicles, propulsion test vehicles, and control test vehicles. This approach demonstrated that early sub-system testing in one phase, if done sufficiently, may be used as the sole basis for producing EDMs in a subsequent phase.

Another approach uses DT to mature the design in an evolutionary fashion with a long-term (6-12 months) assessment in a shipboard environment. The T&E of Navy hull/mechanical/electrical equipment frequently includes long-term operation aboard a Fleet ship. The Reverse Osmosis Desalination Plant (RODP) program is one such program. A unique aspect of the RODP test program was that concurrent at-sea DTs were conducted, and compared different pre-production prototype designs to determine which one should proceed into OT. The initial version RODP was designed, built, and tested aboard a destroyer for an 8-month DT. Based on experience gained in operation and maintenance of the plant and earlier RODPs, an improved, simpler, more reliable design was completed by the Navy and was tested for 10-months aboard a different destroyer. The initial unit was the better-performing unit, and was therefore selected for OT, which was conducted for 9 months and was extremely successful.

Much more reflective of today's evolutionary acquisition approach is to conduct DT and OT in a phased manner on EDMs with incremental performance

enhancements. Occasionally, as part of a system upgrade, a particular operational computer program will be singled out for specific engineering and manufacturing development T&E, including TECHEVAL and OPEVAL, and a separate approval decision on Fleet introduction. Such was the case with the MK 48 ADCAP Torpedo Block Upgrade II (BU II). This effort supported evolutionary updates planned for the system. BU II was subjected to a TECHEVAL and OPEVAL because of the significance of the tactical and operational improvements it made to the system. DT testing was conducted in several phases: (1) by the contractor, to evaluate compliance with the performance specifications; (2) by both the contractor and the Navy, at the program's land-based test site at the Naval Underwater Systems Center, Newport, RI, to assess its operation with the Combat Control System; and (3) aboard an active Fleet SSN for assessment of its utility against new threat submarines under operational conditions and tactics. OT consisted of OPTEVFOR's monitoring the DT testing, as well as conducting a dedicated OT Phase exercise at sea.

Another example of this evolutionary testing of systems is the Combat Control System (CCS) Mk 1. This "system" is comprised primarily of software with some minor processing equipment for command and control of attack and ballistic missile submarines. The system continuously undergoes evolutionary changes during its life. The software upgrades are required for the introduction of new weapons capability, such as vertical and horizontally launched TOMAHAWK. Due to the critical nature of the functions this software controls, OPEVAL was conducted for each of the major functional upgrades, even though the basic architecture remained the same. The upgrade was first tested in the lab and then on the simulator. The program was then downloaded to a submarine platform and subsequently tested with the existing ship interfaces. An OPEVAL was conducted using actual missile launches. The ship testing identified minor human engineering issues that were not obvious during laboratory tests and verified that the upgrade had no deleterious effect on the existing baseline functions.

With regard to Ship Program T&E activities between the new MS C and the Full Rate Production decision, the number and extent of DT and OT phases will vary from one program to another, depending on the complexity of the ship's combat and warfare systems and their risks. Title 20, US Code Section 2399 requires completion of IOT&E prior to proceeding beyond LRIP (FRP Decision). For ships, OPEVAL becomes an OT of the production article. The difference in T&E for ship acquisition versus T&E of weapons systems is that an OT&E on the lead ship is done after the decision has been made for the program in terms of follow-ships contract awards and deliveries.

The T&E program history of the ARLEIGH BURKE Destroyer Program (DDG 51 Class) provides an illustration of Ship Acquisition T&E that used a "build a little, test a little" approach. The program transitioned through early land-based combat systems testing (DT&E and OT&E) at the Combat System Engineering Development Site (CSEDS), Moorestown, NJ. Initially, the testing was conducted on just the Anti-Air Warfare (AAW) portion (to support an LRIP for the lead and follow ships). Then incrementally, as the combat system configuration evolved, the combat system was presented for DT&E and OT&E using the CSEDS. Three major DT and OT events were conducted at CSEDS. Once the lead ship became available for DT and OT testing after delivery, numerous DT and OT events were conducted, starting with an OT phase of the entire lead ship (which focused on integrated combat systems performance). Subsequent DT and OT phases were conducted, addressing the propulsion plant/electrical systems. FOT&E was conducted to address Combat Systems upgrades, starting with Flight II ships and continuing through Flight IIA.

Another ship example involved the Strategic Sealift Ship program. While this program did not involve significant technical risks, there were a number of management challenges in the Navy developing a ship where the Army was to be "user" and Military Sealift Command (MSC) was the "operator." Early in design phases, the program was designated Multi-Service OT&E (MOT&E), with the Navy serving as lead T&E agency and the Army OTA serving as the main

player. OPTEVFOR led the test team, which included membership from Army Operational test agencies. To facilitate planning, the PM held periodic Test Planning Working Group (TPWG) meetings starting nearly 1½ years prior to the OT. The program was also under OSD/DOT&E oversight. The first ship was an early conversion ship, and underwent OPEVAL that conducted on-load and offload of a wide variety of Army vehicles. A "Mission Critical Performance Verification Test (MCPVT)" event, involving on-load and offload of numerous Army vehicles, was conducted to prepare for OT that involved the Navy PM, MSC, Army operational units, and shipbuilder during shipyard testing.

Some programs rely on modeling and simulation (M&S) results to gain approval for limited rate initial production. The AN/SPY-1(D)V Radar Upgrade testing conducted at the AEGIS Combat Systems Development Site (CSEDS), Moorestown, NJ, using five simulations to thoroughly explore the upgrade's capabilities, supported an LRIP buy of 15 (out of a total inventory of 21). This T&E structure was necessary to support the DDG 51 shipbuilding schedule. The shipboard TECHEVAL and OPEVAL would not take place until four years later. The five simulations were developed over a span of two years and accredited by COMOPTEVFOR. In addition to measuring radar performance against the simulation, manned aircraft were used to provide empirical test data to further support the accreditation effort.

5.8 Milestone C

Milestone C is the new Production Readiness Review point. Engineering, DT and OT data to date is reviewed to assess the maturity of the system in support of a Low-Rate Initial Production (LRIP). The amount of testing that supports Milestone C will be agreed to by all parties ahead of time and agreed to in the TEMP. LRIP articles produced as a result of MS C support the T&E process by providing additional test articles used during TECHEVAL and OPEVAL.

5.9 Shipboard Installation

The EDM installation for large systems (sonar, radar, combat system) requires substantial changes to the ship. For smaller programs, a temporary installation is

much easier. In either case, after OPEVAL, the system is removed and the ship is restored to its original condition. However, program managers should be aware that many of the problems encountered during past TECHEVALs and OPEVALs were due to the temporary or artificial nature of the installation. While the program manager can discount such problems because they are not expected to reappear in the production installations, they nevertheless reduce the demonstrated reliability, detract from the perceived system capabilities, and jeopardize the accumulation of sufficient data to support approval for full-rate production. The TECHEVAL and OPEVAL system, although conducted with a prototype, should be very representative of the planned production configuration, including spare parts support and preliminary technical manuals. Also, the contractor or Navy field activity should have additional spare parts available, in case of unanticipated failures in testing. Since a production line may not exist, such parts must be planned well in advance to ensure their availability.

The installation itself should be similar to the expected production installation. Some shipboard electric power conversion equipment failed OPEVAL because of problems caused by two conditions that would not have been present in the production configuration: (1) as a cost savings measure, the equipment was installed in the helicopter hangar (rather than in the usual below-deck compartment), where unanticipated structural interference caused numerous interruptions; and (2) a water-cooled heat exchanger, installed above the electronics cabinet (in violation of standard shipboard practice) leaked, causing the system to be put in deficiency status for two months. OPTEVFOR withheld a recommendation for Fleet introduction because the required reliability and availability had not been demonstrated. The heat exchanger leak highlights another cautionary note. When a system is new, its installation documentation is often incomplete, untried, and seldom covers the unusual installations conditions that can occur for an OPEVAL.

The Plastics Waste Processor was an ACAT IVT program on accelerated development to provide ships an efficient way

to process plastic waste to meet stringent Environmental regulations. The system under test was installed in an aircraft carrier with a general arrangement configuration that did not reflect the final production installation configuration. This was done because the existing equipment could not be taken off-line for the test. The room was quite small, which made access to maintenance panels difficult. The deployment took longer than had been envisioned, requiring maintenance action. The installation made it difficult to perform the maintenance, and the system was faulted in the operational suitability area for not meeting its maintainability threshold and for safety deficiencies. Since operational effectiveness was found to be satisfactory, and there was strong Congressional language to expedite the installations, the OPEVAL report results justified proceeding into production. The PM was required to conduct an FOT&E on a different ship application to demonstrate the fixes to the suitability problems cited in OPEVAL.

5.10 TECHEVAL

TECHEVAL has two purposes: (1) to verify that the system design planned for production meets technical and operational performance requirements; and (2) to verify that the system is ready for OPEVAL. As stated previously, TECHEVAL is generally the last DT&E event of the development. TECHEVAL should be a dress rehearsal of OPEVAL using the same ship, personnel, equipment, and operational environment as that planned for OPEVAL. A Test Readiness Review (TRR) assessment is usually conducted by the T&E IPT immediately prior to major events, such as TECHEVAL, because most current events are too complex and require many too resources to begin without a good scrub of everything that will be done. The TRR is used as a final control point to ensure that everything and everyone is ready to go. A good T&E manager will establish specific exit criteria (performance, personnel, site readiness) that are tracked during the T&E preparation to use as a benchmark. Without such exit criteria, the PM will have no way of knowing when he is ready for test. A key tenet of successful T&E is that you don't go to test unless you are ready for it and there is a good chance your objectives will be met. Without this premise, you would be testing for test's sake.

TECHEVAL should be a dress rehearsal of OPEVAL using the same ship, personnel, equipment, and operational environment as that planned for OPEVAL.

5.10.1 Scope of TECHEVAL

TECHEVAL should be a rehearsal for OPEVAL. It should include T&E of the technical and operational performance, the interoperability of the system with other systems with which it must interface, its compatibility with its environment (e.g., the shipboard operating environment), electromagnetic compatibility, survivability and vulnerability, reliability, maintainability, availability, spare parts support, operator and maintenance technical manuals, training, safety, human factors, and transportability.

The program manager should try to ensure that the OPEVAL crew will be available during TECHEVAL, and that they operate and maintain the system. Not only does this provide an opportunity to prepare them for OPEVAL, it also allows a meaningful T&E of the ILS elements. For example, the preliminary operator and maintenance manuals, support and test equipment, and preliminary Maintenance Requirements Cards (MRCs) should be used by the sailors. The formal maintainability demonstration, whether conducted on the ship or in the factory, should be done using actual fleet personnel. This will provide an early identification of discrepancies, some of which may be correctable prior to OPEVAL. Most program managers are now planning TECHEVAL in two distinct stages: (1) with heavy contractor and field activity assistance to groom the equipment to peak operating conditions, and (2) with almost full operation and maintenance by the crew to prepare them to manage whatever occurs during OPEVAL.

One other important way to enhance traceability from TECHEVAL to OPEVAL is to work with the Operational Test Director (OTD) to use the same test

instrumentation during both exercises. As simple as it seems, this is often overlooked.

The program manager must not be afraid to stress the system during TECHEVAL. OPTEVFOR is sure to do so during OPEVAL. The program manager should be prudent, however, and plan to stress the system only after sufficient testing has occurred so that if the system does fail under stress or over—stress, the resources and time invested in setting up TECHEVAL are not wasted. A failure at the end of TECHEVAL may only require the system be overhauled prior to OPEVAL. A failure early in TECHEVAL may require that the system be shut down, pending repair. This would cause the cancellation and subsequent rescheduling of TECHEVAL, the loss of a great deal of time, and the loss of fleet resources.

The number and type of ship riders should be limited when conducting tests or data collection with fleet forces. Scientists tend to not associate well with the ship's force, and that can lead to a feeling of animosity toward them. Project engineers who know the system, are not afraid to get their hands dirty, and treat the crew as equals, are a better choice. Sending people who do not have specific, full-time jobs should be avoided. The number of high-level personnel should also be kept to a minimum, as the crew will react differently and may not take risks when exercising the system if they feel that they, and not the system, are being assessed.

Some managers have underestimated the importance of the TECHEVAL and the need for testing at-sea. There is heavy emphasis today on using land-based testing, as well as Modeling and Simulation (M&S), to characterize system performance to a much higher degree than at-sea testing. While this makes good economic sense in the short run, it is no substitute for operation by real sailors, on the ship, at sea. Although a system may appear stable in the lab, there is no guarantee it will be that way at sea. Models are only as good as the assumptions that went into them, and land-based test sites are only as good as the interfacing equipment they contain. Only at sea can a system experience the “unknowns”.

The Advanced Combat Direction System (ACDS) TECHEVAL and OPEVAL present a case in point. A majority of the TECHEVAL was performed at a land-based test site. A much smaller portion of TECHEVAL was to be spent at-sea with reduced scope. The land-based testing showed that all thresholds would be easily met. The at-sea portion was cut back to recover a schedule slip. The OT, surprisingly, found ACDS to have numerous deficiencies, discovered when the system was in real-world operation. The simulation used in the land-based testing did not stress the system enough and gave the PM a false sense of ACDS' true performance.

5.10.2 TECHEVAL Plan

The importance and visibility of TECHEVAL, as well as the use of scarce resources involved, demand a well-written test plan and test report. Chapter 3 covers the need for T&E documentation in depth. Several NAVSEA systems have failed OPEVAL due to latent technical deficiencies that were not uncovered in TECHEVAL because of limitations in scope. The plan should ensure the grooming and testing of interfacing systems. Systems that the program manager considers outside his/her responsibility may well be found by OPTEVFOR to be a major consideration in the assessment of the system's overall effectiveness and suitability for Fleet deployment.

5.10.3 TECHEVAL Report

The results of TECHEVAL are used by the OT Readiness Review in judging whether the system is ready for OPEVAL. The program manager must, therefore, plan to have at least a quick-look report available in time to support the review, usually at least two weeks prior to OPEVAL start. The report (both quick-look and final) must describe:

- The scope of testing, including a description of the major test events.
- Equipment configuration highlighting, in particular, known differences between the TECHEVAL, OPEVAL, and planned production versions.

- Limitations to testing that the readers need to be aware of when considering test results and the reports conclusions.
- Test results, both quantitative and qualitative.
- Comparison of test results to TEMP requirements.
- Any additional considerations that would be useful in the OT readiness review.

Copies of the TECHEVAL report must be sent to OPNAV and OPTEVFOR for information, in addition to those seen at the OT Readiness Review. A full discussion on test reporting is found in Chapter 3.

5.11 OT Readiness

SECNAVINST 5000.2B establishes the requirements for SYSCOM/PEO/DRPM certification of readiness of each system to enter major OT&E events. The readiness for OT is not a one-time event, but rather, a continuum of planning and assessment throughout the development that culminates in one certification of readiness. The certification is made to COMOPTEVFOR, usually by Naval message, with information copies sent to other interested commands. Where a waiver is required from one or more of the CNO Certification Criteria, the request is sent to CNO N912. It is OPNAV and headquarters policy, based on many lessons learned, not to enter OPEVAL until the readiness criteria have been satisfied and it is believed the system will pass OPEVAL; i.e., obtain a positive recommendation from OPTEVFOR for Fleet introduction of the system. The OPEVAL certification criteria is applied rather rigorously in the case of OPEVALs or FOT&Es and for combined DT/OTs when the data will be used by OPTEVFOR to help resolve future COIs. For OAs, instead of the more formal and rigorous "Certification of Readiness" for OPEVAL, a notification to COMOPTEVFOR, sometimes called a "Notification of Availability for Testing," provides an affirmation that the system is safe to operate and that minimal logistic support exists. The Notification does not address whether the system meets the technical and operational requirements. Rather, the Notification is aimed at getting quickly into operational assessment/testing

without the delays involved in formal certification procedures.

It is Navy policy -- and it is a programmatic practice strongly embraced by NAVSEA -- that a system shall not proceed to OPEVAL until there is little risk that it will not pass that OPEVAL, i.e., that it will be not found by COMOPTEVFOR to be operationally effective and operationally suitable.

5.11.1 Readiness Criteria

The following are the SECNAVINST 5000.2 criteria that the Navy has identified to be used in assessing readiness. Each should be briefly addressed during the OT Readiness Review:

1. The TEMP is current and approved.
2. All TEMP-specified DT&E objectives and performance thresholds have been met, or are projected to be met at system maturity. All TEMP-specified DT&E tests have been completed and the reports (including TECHEVAL report) are published. The results of DT&E demonstrate that:

- (A) Engineering is complete,
- (B) All DT&E objectives and

performance thresholds have been met, or projected to be met, at system maturity. The results indicate that the system will perform successfully in OT&E and will meet the criteria for approval at the next program decision milestone (e.g., full-rate production on completion of OPEVAL). All DT&E testing data has been published and distributed. With the exception of combined DT/OT, the DT test reports are distributed 30 days prior to the start of operational testing.

3. The results of DT&E (and previous OT&E) demonstrate that all significant design problems (including compatibility, electromagnetic environmental effects, interoperability, survivability/vulnerability, reliability, maintainability, availability, human factors, system safety, and logistics

supportability) have been identified and corrective actions are in process.

4. System operating and maintenance documents, including 3-M and preliminary allowance parts list (PAPL), have been distributed to COMOPTEVFOR.
5. Adequate logistic support, including spares and repair parts, support/ground support equipment, etc., is available as documented in the TEMP and integrated logistics support plan (ILSP).
6. The applicable system technical documentation, such as failure mode effect and criticality analyses, level of repair analyses (LORA), life cycle cost (LCC), and logistic support analyses (LSA) have been provided to COMOPTEVFOR.
7. The OT&E manning of the system is the same (in numbers, rates, ratings, and experience level) as is planned for fleet units under normal operating conditions.
8. The Navy training plan has been approved and provided to COMOPTEVFOR.
9. Training for personnel who will operate and maintain the system during OT&E (including OPTEVFOR personnel) has been completed, and this training is representative of that planned for fleet units under the Navy training plan.
10. All resources required for operational testing (instrumentation, simulators, targets, expendables) have been identified and all appropriate documents are available.
11. The system provided for OPEVAL/FOT&E, including software and the total logistics support system, is production-representative.
12. All threat information required for OPEVAL/FOT&E (i.e., threat system characteristics and performance, electronic countermeasures, force levels, scenarios and tactics) is available and a list of such information (including security classifications) has been provided to COMOPTEVFOR.
13. The system safety program has been satisfactorily completed.

14. The system complies with Navy occupational safety and health/hazardous waste requirements where applicable.

15. Software maturity metrics analysis demonstrates that the software is stable and expected to perform at a level commensurate with the operational test phase.

16. For software qualification testing (SQT), a statement of functionality, describing the software capability, has been provided to OPTEVFOR.

17. For programs employing software, there are no unresolved Priority 1 or 2 Problem Reports (SPR), and all Priority 3 problems are documented with appropriate impact analysis.

A detailed checklist based on these criteria is provided in Appendix A. This checklist was prepared by the NAVSEA T&E Office based on experience in past readiness reviews.

5.11.2 Getting Ready For OT

Figure 5-2 shows a notional timeline for getting ready for OT. At six to eight weeks prior to the planned certification date, the T&E IPT refocuses their meetings specifically toward the test results and readiness items required necessary to support the Operational Test Readiness Review (OTRR). This emphasis allows for time to define, address, and correct potential certification issues. Approximately one month prior to the planned certification, a memorandum convening the review is released to OTRR participants. The preferred and usual method of completing the review is with a formal meeting of all

participants. However, the DEPCOM/PEO may determine that a formal meeting is not necessary and elect to certify using a different method. No sooner than one week prior to the certification, a pre-review is conducted at the working level with the applicable documentation to evaluate the readiness of the system. This pre-review will outline any concerns so that the program manager will be better prepared for the OTRR. The documents normally supporting these reviews include:

- (1) The proposed certification message, addressing the criteria contained in reference (a).
- (2) The system TEMP.
- (3) The requirements document (e.g., the Operational Requirements Document (ORD))
- (4) TECHEVAL/DT III plan. (NOTE: Such exercises must have, as stated objectives, not only the demonstration of design requirements, but also the demonstration of readiness for the OPEVAL/FOT&E.)
- (5) DT&E report(s).
- (6) OPEVAL/FOT&E plan.
- (7) Previous OT&E report(s).
- (8) Integrated Logistics Support (ILS) Plan.
- (9) Navy Training Plan.
- (10) Technical manuals (samples of operators, maintenance, and Illustrated Parts Breakdown manuals).
- (11) Planned Maintenance System (PMS) documentation.
- (12) Allowance Parts List (APL) (preliminary).
- (13) Reliability and Maintainability plans and reports.
- (14) System safety and Environmental reports.

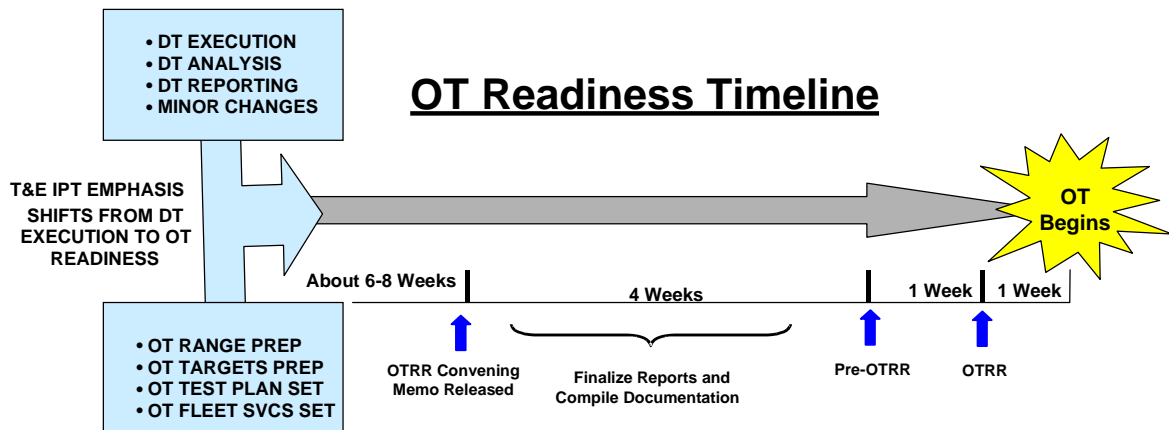


Figure 5-2 – OT READINESS TIMELINE

5.11.3 OT Readiness Review

The process of assessing system readiness to commence OPEVAL and FOT&E exercises is to conduct an Operational Test Readiness Review (OTRR). The Deputy Commander, Assistant Deputy Commander (SES or Flag level) or PEO responsible for the acquisition program, chairs this review. Participants are limited to senior persons authorized to take a position in the respective functional area and make recommendations to the chair. The OTRR is the DA's certification of readiness. As such, the risk is on the developers, though there may be heavy influence by outside agents to proceed into testing. Participants generally include:

Members:

- (1) Program Manager
- (2) Program T&E Manager
- (3) Directorate/PEO Staff as appropriate
- (4) Test and Evaluation Office focal
- (5) Program Engineering Manager
- (6) Lead Field Activity Representative
- (7) Program Integrated Logistics Support Manager

Advisors:

- (1) CNO N912 and COMOPTEVFOR OTD Representatives
- (2) CNO Sponsor for the program
- (3) OSD T&E action officers (DOT&E and DT&E)

- (4) Ship/Squadron representatives (if necessary)

The first seven comprise the advisory body of developers who are the principals. The remaining attendees are interested parties from whom the certification authority solicits comments.

A Certification of OT Readiness Checklist (Attachment A), detailing the OPNAV Certification criteria, is used as a guide to ensure a complete and thorough review.

5.11.4 OTRR Brief

The OTRR Brief is a presentation by the program manager addressing the topics contained in the certification message. The program office presents performance, testing, and test participant readiness results and risk areas, as well as corrective actions for any deficiencies. Personnel from the CNO program sponsors, T&E coordinator's offices, OPTEVFOR Operational Test Director, and other empowered technical personnel provide input and advice. After the review, the DEPCOM/PEO makes the decision to proceed with the OT and signs the certification message.

The overall agenda of the OTRR and of the presentation is as follows:

Description/Overview

System
Schedule/T&E Funding

Documentation Status (TEMP, ORD, DT/OT Test Plan, WSESRB, etc.)

DT Testing/TECHEVAL

Description of Tests

Test Results (Compared to DT and OT TEMP Criteria)

Key Technical Problems/Status

Computer Software Maturity (PTR, Longevity, Scope/Breadth of Testing)

TECHEVAL/OPEVAL Configuration Differences

OT Readiness

Plan/ Schedule/Scope of Testing

Equipment readiness

Test platform Readiness

Logistics

Safety/Environment

OPSEC (if applicable)

Conclusions

Recommendations

5.11.5 Certification of Readiness

As outlined in SECNAVINST 5000.2B, the certification is made to COMOPTEVFOR (with copy to the CNO) by a message if there are no waivers, or to CNO N912 (copy to COMOPTEVFOR) if waivers are required. Waivers only allow the OT to proceed and are not waivers from meeting specific performance objectives. Waivers will be requested for minor items only, provided there is appropriate justification. Prior to the start of the OPEVAL, the program manager must correct the deficiencies that would jeopardize successful completion of the exercise. If such corrections cannot be made prior to the usual release date for the certification message, but DEPCOM/PEO is confident that the correction will be properly made prior to the scheduled commencement of OPEVAL, the message shall identify the status as such.

Certification waivers only allow the OT to proceed and are not waivers from meeting specific performance objectives.

5.11.6 De-certification/Re-certification

A system undergoing OPEVAL or FOT&E can be placed in deficiency status if it has been determined that some problem exists which would prevent completion of the OT. This was the case with the MK50 Torpedo where it was discovered that most of the units prepared for OT contained a manufacturing lot defect that would cause the units to run poorly. The torpedo was De-certified by the PM upon discovery. In another case, OPTEVFOR identified a series of new reliability problems on the AN/SQQ-32 Sonar System that essentially meant that the OT could not be conducted. Thus the system was de-certified to conserve test assets. A de-certification, which essentially ends the OT event, can be initiated by the PM or COMOPTEVFOR. If OPTEVFOR issues a message placing the system in deficiency status, a full re-certification review must be conducted to certify the system's readiness to recommence the exercise in accordance with SECNAVINST 5000.2B.

5.12 OPEVAL

The OPEVAL is an independent operational test conducted by OPTEVFOR on the system in its intended operating environment with fleet operators. OPTEVFOR writes the plan, obtains the necessary Fleet resources, directs the test, analyzes the data, and writes the report. Far more than in any other OT&E exercise in the program, OPTEVFOR rigidly insists on complete independence from the development team in the OPEVAL. As such, the PM must ensure that development testing demonstrates that the system operates using fleet crew without the need for DT personnel involvement. In addition, the PM must ensure that everything is in place to execute the test with limited/or no involvement by the DT community. This means that the PM should clearly understand how OPTEVFOR will operate and assess the system in OPEVAL. This can be ensured with a close working relationship with the OTD and ensuring that all test plans are coordinated. Remember, the PM cannot certify that the program is ready for OPEVAL and that there is a good chance of passing it, unless the DT program reflects what will be done in OT. The independence of the OT process can be

maintained, while maintaining good communications with your OTD.

5.12.1 OPEVAL Test Plan

OPTEVFOR writes the plan to fully exercise the system in its intended operational environment. The tests are structured to gather sufficient data to make an assessment of its operational effectiveness (E-tests) and its operational suitability (S-tests), and to make a recommendation to the CNO about approving the system for production and Fleet introduction. Frequently, the plan, or a draft of the plan, is received 30 days prior to the start of OPEVAL. The program manager is not part of the OPEVAL Plan approval process. Therefore, if the program manager has an issue with anything in the OPEVAL Test Plan, he needs to quickly rectify the issue with OPTEVFOR prior to the start of OPEVAL. If reconciliation is not possible, the program manager must determine the impact of the issue and clearly spell out the situation in the certification for OPEVAL message.

In one program, there was a significant misunderstanding between the OTD and the program manager on the emplacement procedures for a series of underwater fuses that a review of the plan would have uncovered. The OTD, newly reported aboard at OPTEVFOR, used procedures that varied from those called for in the design. The OTD described his intentions in the plan, but was unaware of the disconnect until he received the system technical manual several days before the commencement of OPEVAL. When he received the manual, he decided to follow the plan anyway because of constraints that the NAVSEA emplacement procedures would have imposed on the Fleet. Few of the units operated properly and the system failed OPEVAL. The OPEVAL Test Plan must be recognized as one of many vehicles, and a critical one, for the program manager to assess OPTEVFOR's understanding of the system requirements. The SECNAVINST on T&E (5000.2) requires the plan to be published 30 days prior to testing. NAVSEA includes a review of the plan as part of the OT Readiness Review prior to certification. The onus is on the program manager to work closely with the OTD on the plan.

5.12.2 Resources

OPTEVFOR obtains the Fleet resources (e.g., ships, aircraft, range time, targets, expendable ordnance) for the OPEVAL. However, any funding required must be provided by the program manager. Typical expenses include range support personnel, expendable ordnance target presentation costs, and data analysis, if OPTEVFOR must have it done by an outside activity. More about funding T&E is covered in Chapter 2. However, the program manager should be aware that OPEVAL, in particular, can be significantly expensive. One of the purposes of the TEMP is to force the program manager and OPTEVFOR to identify (and perhaps negotiate, if necessary) early in the program what resources will be required, so that they can be routinely requested in the Planning, Programming and Budgeting System (PPBS) cycle.

5.12.3 Conduct of OPEVAL

OPTEVFOR insists that for an OPEVAL, no program personnel, including Navy field activity or contractor personnel, be present, primarily to ensure that the integrity of normal operating Fleet conditions is maintained during the testing without extraordinary assistance. This does not mean that DT representatives cannot support the OT, but that the roles and actions must be clearly delineated in a Memorandum of Agreement. These can be for items such as preparing targets or maintaining interfacing systems not directly coupled to the system under test. OPTEVFOR attempts to estimate the full spectrum of a system's capability, and the limitations that can be expected, once the production units are deployed. OPTEVFOR evaluates the performance of the production-representative model used in OPEVAL against the most demanding operational requirements. Little allowance is made for the fact that it is only a prototype. The obligation is placed on the Systems Command to ensure that the prototype and its installation are representative of what is planned for production. To the extent that they are not representative, OPTEVFOR withholds judgment on the system's acceptability. OPTEVFOR generally makes favorable fleet introduction recommendations only on what has been proven, and will not extrapolate OPEVAL data to cover mission scenarios or

system variants not actually tested. One-time installation problems, unanticipated and uncontrollable limitations to the scope of testing, and deficiencies caused by system components not planned for the production version are frequently not discounted when OPTEVFOR draws conclusions from the OPEVAL results and makes recommendations to OPNAV. This is why the entire engineering and manufacturing development T&E effort, and the TECHEVAL in particular, should contribute to ensuring that the system will perform successfully in OPEVAL.

5.12.4 OPEVAL Report

OPTEVFOR's goal is to issue the report within 90 days after project operations. If, for programmatic reasons, a quick-look report is needed in a shorter amount of time, the CNO must specifically direct OPTEVFOR to provide it. However, OPNAV and OPTEVFOR resist such reports. When OPTEVFOR does issue such a report, he usually includes a caveat that the conclusions are preliminary and may be revised when analysis is complete. When a Quick-look report is desired, the PM must request so from the CNO in the OT Certification message. Realize too, that when OPTEVFOR is required to issue a quick-look report, it may delay issuance of the final report.

5.13 Full-Rate Production (FRP) Review

At the FRP review, a decision is made to enter full-rate production, to continue low-rate initial production, or not to enter production at that time (this latter decision can lead to a cancellation of the program or to additional system development and T&E). The T&E-related questions that should be answered at FRP are:

- Has all TEMP-required T&E been completed?
- Have all T&E objectives and performance thresholds been met?
- Are there a plan of action and milestones to correct outstanding discrepancies?

For programs under NAVSEA's cognizance, the NAVSEA T&E Office reviews these items prior to the FRP review. As an advisor, the NAVSEA T&E Office reports the results of his review to the Review Chairman, with a recommendation on whether corporate NAVSEA should support the request for production that the program manager plans to make at the decision meeting. The results are documented in the final decision memorandum.

CHAPTER SIX FULL-RATE PRODUCTION AND DEPLOYMENT

6.0 INTRODUCTION

The transition from System Development and Demonstration and initial production to full-rate production and deployment is accomplished by developing the engineering model design and applying it to production hardware for delivery to the Fleet. The objectives of the production effort are to: (1) achieve production of authorized quantities, on schedule and within budget; and (2) achieve readiness for system deployment. The objectives of the deployment effort are to achieve a high level of operational readiness for the deployed system; i.e., personnel assignment, training, maintenance, supply/spare support, and overhaul, alteration and repair. The Full Rate Production (FRP) and Deployment decision is the major decision point where production of systems for permanent installation on Fleet units, land-based configuration and training facilities, or for inventory is authorized. The Rate Production and Deployment phase includes (as required) DT-III and OT-III conducted on production items.

6.1 Developmental T&E DT-III

DT-III is the developmental test and evaluation conducted after the full-rate production and deployment decision has been made. The use of performance specifications for production poses special T&E considerations. Additional testing may have to be planned to revalidate performance of the production system that otherwise would not have been needed if the identical system developed was procured. Special emphasis must be placed on ensuring performance in environmental compliance, interoperability, and software compatibility. DT-III is usually conducted:

(1) To verify the effectiveness of product improvements or corrections made after TECHEVAL and OPEVAL.

(2) To demonstrate the adequacy of production line redesigns as a result of production changes or early follow-on operational test and evaluation (FOT&E).

(3) To demonstrate readiness for later FOT&E exercises.

“The use of performance specifications for production poses special T&E considerations”

6.2 Follow-on Operational T&E

OPTEVFOR is assigned to conduct an OT&E exercise after the FRP decision to evaluate the correction of deficiencies identified during TECHEVAL and OPEVAL, as well as to conduct operational testing not conducted prior to FRP. Standard practice is to include provisions for Follow-on Operational Test and Evaluation (FOT&E) exercises in early issues of the TEMP, even though the specific scope and objectives cannot be delineated meaningfully until late in the System Development and Demonstration.

FOT&E is OT-III, conducted to verify corrections made as a result of TECHEVAL and OPEVAL, and to complete testing deferred from IOT&E. OT-III is conducted on production hardware to verify correction of deficiencies of production systems, testing of the system in new environments, or experimenting with new tactical applications or against new threats. The program manager usually funds OT-III with production funds if addressing deficiencies, or with R&D funds if examining new capabilities.

6.2.1 Certification of Readiness for FOT&E

When the objectives of an FOT&E exercise are to conduct testing deferred from OPEVAL, or to demonstrate the correction of deficiencies identified in OPEVAL, the DEPSYSCOM/PEO must certify to the CNO and COMOPTEVFOR that the system is ready to commence the exercise. The criteria and procedures for certification are the same as those for OPEVAL, described in Chapter Five. Program managers should plan for such an exercise in much the same way as they plan for in OPEVAL, including the conduct of the preparatory DT-III exercise as a rehearsal for the FOT&E.

6.3 Technology Insertion and OT

There are many instances in which upgrades are introduced into production systems. The reasons could include the need to take advantage of emerging technology for new warfighting methods, replacement of technology to lower total ownership cost, replacement of obsolescent technology as a matter of necessity, or

to facilitate the introduction of some other interfacing system. In any case, a decision must be made on how much testing is required, and on what level of formality is necessary. The SECNAVINST 5000.2B, Section 1.4.5, provides a decision matrix for assessing when a modification warrants a formal testing and a TEMP. Whether formal testing is required under an "ACAT" program or as part of a contract requirement should not influence the requirement for testing upgrades to a degree that has quantified the risks to the fleet introduction authority. Just because a modification or change may not formally require a TEMP does not mean testing is not required.

6.4 Production Acceptance Test and Evaluation (PAT&E)

Production Acceptance Test and Evaluation (PAT&E) is defined as that testing conducted on production items to demonstrate that systems meet contract specifications and requirements. PAT&E also includes the testing necessary to demonstrate that items/systems are properly installed and operable on a ship or aircraft platform. Most PAT&E, including funding, is the responsibility of the program manager.

6.4.1 PAT&E for Control in Production

For factory acceptance, a distinction must first be made between acceptance testing of a "first article" and testing of the "follow-on buys". First article factory acceptance, environmental and reliability tests, and maintainability demonstrations may not require retest if already accomplished prior to production acceptance. Tests that may be necessary for production acceptance are as follows:

- Manufacturing screening tests.
- Burn-in tests.
- Random sampling tests.
- In-process tests.
- Environmental tests.
 - Shock.
 - Vibration.
 - Temperature sequence.
 - Humidity.
 - Salt fog.
 - Inclination.
 - Magnetic field environment.
 - Accelerated life.
- R&M demonstration tests.
- Proof tests.
- Electromagnetic Interference tests.

6.4.1.1 Manufacturing Screening Tests

The purpose of screening tests is to identify defective parts and workmanship prior to installation. Careful manufacturing screening tests reduce assembly and subassembly rejects and rework costs. This type of test is generally applied to electronic components and includes such tests as random vibration, thermal cycling, failure-free operation, electronic parts screening, and particle impact noise detection.

6.4.1.2 Burn-in Tests

The burn-in test is designed to stress a device to identify failure-prone components. This type of test usually pertains to electronic devices in which "infant mortality" is a potential problem. The burn-in test is a stress test and does not simulate an operating circuit. The primary intent of this test is to prevent random failure in an installed system after delivery, when the impact of failure can be costly.

6.4.1.3 Random Sampling Tests

This sampling test randomly selects finished items/components to be tested for conformance to specification. The intent of this test is to provide a measure of assurance that all production items are meeting specifications.

6.4.1.4 In-process Tests.

The in-process test is basically a sampling test, similar to the above test, except that it selects samples from some point on the production line. In-process tests minimize failure at higher levels of assembly through early recognition and correction of defects.

6.4.1.5 Environmental Tests

Environmental tests subject the device/system to a realistic environment, to determine if the device/system can withstand specifications established in the production contract. The sea environment is probably the extreme "worst case," making it especially important to test the system against environmental conditions.

6.4.1.6 Reliability and Maintainability Demonstration Tests

Reliability and maintainability demonstrations, if required, are normally conducted in the Engineering and Manufacturing Development phase, but may be deferred to the Production and Deployment phase. Demonstration tests have been repeated when subsequent design changes or unexpected environmental problems were

suspected to have significant impact on reliability and maintainability.

6.4.1.7 Proof Tests

Proof tests are acceptance tests that provide a means of examining production devices/systems under operational conditions. A prime example of proofing is the torpedo proofing test, which includes an in-water test of the complete end-item torpedo under realistic operating conditions. This test is usually invoked as a condition of Government acceptance of the torpedo. Proof testing can be applied to any system, whether expendable or non-expendable. The expense of conducting a proof test on an expendable system must be weighed against the benefit. (Proof testing of expendable items is done by testing a statistically significant number of units to represent the entire production run.)

6.4.1.8 Electromagnetic Interference Tests

EMI testing is conducted to ensure that the system meets its performance requirements in the operational EM environment.

6.5 PAT&E for Control in Installation and Operation

PAT&E test programs must include tests necessary to ensure proper system shipboard installation; e.g., pre-installation test, installation checkout and circuit test, system operability test, intersystem operability tests, and ship acceptance trial tests. These tests are described in NAVSEA 0900-LP-095-2010, Ship Construction Tests and Trials Manual.

6.6 In-service T&E

During the deployment phase, the 3-M (Navy Maintenance and Material Management) system is a primary source of equipment failure information from the Fleet. The 3-M system consists of two subsystems: Planned Maintenance System (PMS), and Maintenance Data System (MDS).

The PMS provides each ship with a simple and standard means for planning, scheduling, controlling, and performing planned maintenance and online tests of all equipment.

The MDS is the means by which maintenance personnel report corrective maintenance on specific categories of equipment, except that submarines report corrective maintenance experience on all equipment. The MDS provides information about certain Fleet

maintenance and maintenance support actions for use by Navy management, with particular emphasis on providing information at the shipboard level.

Numerous reports are available to any command from the 3-M central data bank. These reports yield information concerning equipment maintainability and reliability, equipment alteration status, man-hour expenditures, material usage and costs, and the material condition of the Fleet. The usefulness of the MDS depends upon the accuracy, adequacy, and timeliness of the information reported into the system; it is the system in which potential benefits are directly proportional to the efforts applied. Current programs for improving reliability, maintainability, and logistic support of Fleet equipment rely, therefore, upon conscientious adherence to reporting procedures.

The MDS reporting identifies the deferral or completion of a maintenance action. In some instances, the reports document maintenance actions that are not equipment-related, such as services. The 3-M system includes periodic system checks and inspections to ensure that reliable data are available and that appropriate corrective actions are taken.

6.7 Assessment Continuation

Effective data collection and analysis are essential components of a successful T&E program. Assessment and evaluation of a system are continuing efforts throughout production and deployment. Future T&E efforts depend heavily on "lessons learned" from previous procurement, and ways to improve the T&E process must continually be sought. Thus, effective T&E programs must incorporate continuing evaluation of newly procured systems deployed in their operational environment. Program managers often fail to plan for this, and the resulting problems, as experience has shown, can be serious.

Successful ongoing assessment requires a means to identify, report, and analyze the cause of all failures and provide for appropriate corrective action, such as design changes, reduced stresses, manufacturing process changes, and improved quality control.

Appendix A

CHECKLIST FOR OT READINESS REVIEW

1. THE TEMP IS CURRENT AND APPROVED.

- a. Does the TEMP depict the phase of testing being certified and when was it last approved?
 - (1) Are the TEMP thresholds current and are they consistent with requirements documents?
 - (2) Does the TEMP depict the phase of testing being certified for?

2. ALL DT&E OBJECTIVES AND PERFORMANCE THRESHOLDS HAVE BEEN MET, OR PROJECTED TO BE MET AT SYSTEM MATURITY. THE RESULTS INDICATE THAT THE SYSTEM WILL PERFORM SUCCESSFULLY IN OT&E AND WILL MEET THE CRITERIA FOR APPROVAL AT THE NEXT PROGRAM DECISION MILESTONE (E.G. FULL RATE PRODUCTION ON COMPLETION OF OPEVAL). ALL DT&E TESTING DATA HAS BEEN PUBLISHED AND DISTRIBUTED. WITH THE EXCEPTION OF COMBINED DT/OT, THE DT TEST REPORT ARE DISTRIBUTED 30 DAYS PRIOR TO THE START OF OPERATIONAL TESTING.

- a. Has all TEMP-Specified DT&E testing been completed and reports published?
 - (1) Have the final, signed copies of the DT&E reports been provided for the OT Readiness Review?
 - (2) Have the TECHEVAL quicklook (if applicable) and DT&E final reports been provided to:
OPNAV T&E Coordinator
OPNAV Requirements Officer
COMOPTEVFOR
- b. Have all TEMP-Specified T&E performance thresholds been met?
 - (1) Have all the DT&E thresholds been successfully demonstrated?
 - (2) Have all the OT&E thresholds been successfully demonstrated?
 - (3) For any DT&E deficiencies resulting from the TECHEVAL testing, how are the deficiencies being remedied?
 - (4) Has a 25 hour stress test been successfully completed?
 - (5) What other longevity testing has been completed?
- c. Is there a high probability that the system will perform successfully in OPEVAL/FOT&E and meet the criteria for AFRP (if applicable) on completion of OPEVAL/FOT&E?
 - (1) Has the program manager reviewed OPTEVFOR's test plan in depth?
 - (2) Is the OT analysis methodology consistent with that used in DT?
 - (3) Are there any remaining areas of disagreement?
 - (4) Is there an approved copy of the OPTEVFOR test plan available for the OT Readiness Review Board?

3. THE RESULTS OF DT&E (AND PREVIOUS OT&E) DEMONSTRATE THAT ALL SIGNIFICANT DESIGN PROBLEMS (INCLUDING COMPATABILITY, ELECTROMAGNETIC ENVIRONMENTAL

EFFECTS, INTEROPERABILITY, SURVIVABILITY/VULNERABILITY, RELIABILITY, MAINTAINABILITY, AVAILABILITY, HUMAN FACTORS, SYSTEM SAFETY AND LOGISTICS SUPPORTABILITY) HAVE BEEN IDENTIFIED AND CORRECTIVE ACTIONS ARE IN PROCESS.

- a. Are there any ECP's still outstanding against requirements?
 - (1) What is the impact to the overall system performance?
- b. What testing has been done to demonstrate any late fixes?
 - (1) How do these tests compare with the rigor of TECHEVAL?
 - (2) Have fixes been installed in all test units?
- c. Have the applicable environmental tests been completed?
 - (1) If not, will the OT&E subject the system to environments not tested to?
 - (2) Has OPTEVFOR concurred with the schedule for deferred testing?
- d. Has EMI/EMC testing been done of the system installed in its operational environment?
 - (1) Are there any restrictions placed on the system, or interfacing systems as a result of these tests?
- e. Has the TEMP A_0 threshold been demonstrated?
- f. Has the mission profile within which A_0 is to be measured been identified by the program manager (e.g., in the TEMP)? Is it clearly reflected in the system operating guidelines and technical manuals? Does the OPTEVFOR test plan expand performance beyond that profile?

4. SYSTEM OPERATING AND MAINTENANCE DOCUMENTS, INCLUDING 3-M AND PRELIMINARY ALLOWANCE PARTS LIST (PAPL), HAVE BEEN DISTRIBUTED TO COMOPTEVFOR.

- a. What is the status (accuracy and adequacy) of technical manuals?
- b. Has the tactical guidance document been approved by the appropriate authority and distributed to the platform and OPTEVFOR?
- c. What is the status (accuracy and adequacy) of Preliminary APLs?
- d. What is the status (accuracy and adequacy) of Interim or Preliminary Allowance Equipage Lists (AELs)?
- e. What are the status and quality (clarity, accuracy and adequacy) of Maintenance Requirement Cards (MRCs)?
- f. What was the sailor assessment of:
 - technical manuals?
 - parts lists?
 - PMS Documentation (MRCs)?
 - test equipment?
 - training?
 - safety?
 - other?
- g. Have the deficiencies uncovered during TECHEVAL been corrected in:
 - technical manuals?
 - parts lists?
 - MRCs?
 - training?
 - safety?
 - test equipment?
 - other?

- h. Have the technical manuals been validated and verified with input from the Fleet? If negative, when are they scheduled to be accomplished?
- i. Does the parts list identify the system to the piece part level?
- j. Does the parts list reflect the failure rates and Level of Repair (LOR) in accordance the maintenance concept?

5. ADEQUATE LOGISTIC SUPPORT, INCLUDING SPARES AND REPAIR PARTS, SUPPORT/GROUND SUPPORT EQUIPMENT, ETC., IS AVAILABLE AS DOCUMENTED IN THE TEMP AND INTEGRATED LOGISTICS SUPPORT PLAN (ILSP).

- spares and repair parts
- special tools
- other (trucks, dollies, etc.)

- a. Is this support available for all subsystems and all interfacing systems?
- b. Has someone recently physically checked that the spare parts and Maintenance Assist Modules (MAMs) have been delivered for OPEVAL/FOT&E and are available? When were they last checked?
- c. Does the spares and repair parts listing reflect high failure rate items and critical failure needs?
- d. Do the repair parts have National Stock Numbers whenever possible?

6. THE APPLICABLE SYSTEM TECHNICAL DOCUMENTATION SUCH AS FAILURE MODE EFFECT AND CRITICALITY ANALYSES, LEVEL OF REPAIR ANALYSES (LORA), LIFE CYCLE COST (LCC), AND LOGISTIC SUPPORT ANALYSES (LSA) HAVE BEEN PROVIDED TO COMOPTEVFOR.

- Integrated Logistics Support Plan (ILSP) (Production and Deployment Phase)
- Failure Mode and Effects Analysis (FMEA)
- Level of Repair Analysis (LORA)
- Life Cycle Cost (LCC)
- Logistic Support Analysis (LSA)
- other necessary supporting documentation

- a. Is the ILSP (P&D) adequate for OPEVAL use?

7. THE OT&E MANNING OF THE SYSTEM IS THE SAME (IN NUMBERS, RATES, RATINGS, AND EXPERIENCE LEVEL) AS IS PLANNED FOR FLEET UNITS UNDER NORMAL OPERATING CONDITIONS.

- in numbers?
- in rates?
- in rating?
- in experience level?

- a. If the manning is not the same, what is the impact?
- b. Is there any impact of maintenance requirements on ship's manning?

8. THE NAVY TRAINING PLAN HAS BEEN APPROVED AND PROVIDED TO COMOPTEVFOR.

- a. Has the approved NTP been provided to COMOPTEVFOR?
- b. If the NTP is not approved, has a waiver to proceed to OPEVAL without it been granted by OPNAV?
- c. Is the planned training adequate for:
 - operators?
 - maintenance personnel?
- d. What is OPTEVFOR's preliminary assessment of the NTP?

- e. What is the ship crew's preliminary assessment of the NTP?

9. TRAINING FOR PERSONNEL WHO WILL OPERATE AND MAINTAIN THE SYSTEM DURING OT&E (INCLUDING OPTEVFOR PERSONNEL) HAS BEEN COMPLETED, AND THIS TRAINING IS REPRESENTATIVE OF THAT PLANNED FOR FLEET UNITS UNDER THE NAVY TRAINING PLAN.

- install?
- operate?
- maintain?

- a. Have the OPTEVFOR personnel received the same training?
- b. If the training has not been completed, when will it be completed?
- c. Where was the OPEVAL/FOT&E crew obtained?
- d. How much training did each of the members of the OPEVAL/FOT&E crew receive?
- classroom
 - hands-on
 - maintenance
 - operations
 - factory
 - Land Based Test Site (LBTS)
- e. Does the training compare favorably with that planned for Fleet units?
- f. Have back-up personnel been trained in case planned operators cannot perform or in case they are transferred during OPEVAL?
- g. What is the attitude of the operators toward the training?
- h. What is the attitude of the operator toward the planned OPEVAL/FOT&E?

10. ALL RESOURCES REQUIRED FOR OPERATIONAL TESTING (INSTRUMENTATION, SIMULATORS, TARGETS, EXPENDABLES) HAVE BEEN IDENTIFIED AND ALL APPROPRIATE DOCUMENTS ARE AVAILABLE.

- special health/hazardous waste teams or equipment?
 - targets (required/expended)?
 - special instrumentation?
 - support equipment?
 - installation removal requirements?
 - expendables?
 - spare and repair parts?
 - maintenance assist modules (MAMs)?
 - personnel?
 - test site(s)?
 - test range(s)?
 - aircraft?
 - ship(s)/boats?
 - submarine(s)?
 - supporting systems (including computers/software)?
 - models and simulators?
 - personnel training (officers/enlisted/civilians)?
 - planned travel?
 - operational security (OPSEC)?
 - special tools?
 - test equipment required for maintenance?
 - other?
- a. Is instrumentation to be used for this OT&E adequately calibrated and adequate to meet OT&E accuracy requirements?

- b. Has the instrumentation package been validated?
- c. Are the following resources the same as used for TECHEVAL?
 - targets, all types, variants, and augmentations
 - special instrumentation
 - support equipment
 - expendables
 - spare and repair parts
 - maintenance assist modules (MAMs)
 - personnel
 - test site(s)
 - test range(s)
 - aircraft types
 - ship(s)/boats
 - submarine(s)
 - supporting systems (including computers/software)
 - personnel training (officers/enlisted/civilians)
 - security
 - special tools
 - other
- d. Are there available back-ups for critical resources?

11. THE SYSTEM PROVIDED FOR OPEVAL/FOT&E, INCLUDING SOFTWARE AND THE TOTAL LOGISTICS SUPPORT SYSTEM, IS PRODUCTION REPRESENTATIVE.

- a. Are the known differences between this hardware and software and the production hardware and software understood by OPTEVFOR?
- b. Could any of the known configuration differences cause OPTEVFOR to limit the scope of testing during OT and is OPTEVFOR aware of the impact of these differences?
- c. Has the software been exercised and stressed in production representative hardware to ensure it correctly performs its intended function (is technically ready for fleet release), and thoroughly tested and validated to meet CNO thresholds and performance requirements.[NOTE: IF THIS IS NOT THE CASE, SPECIFY IN DETAIL THE DIFFERENCE BETWEEN THE SYSTEM TO BE USED FOR TEST AND THE FINAL PRODUCTION CONFIGURATION]

12. ALL THREAT INFORMATION REQUIRED FOR OPEVAL/FOT&E (I.E., THREAT SYSTEM CHARACTERISTICS AND PERFORMANCE, ELECTRONIC COUNTERMEASURES, FORCE LEVELS, SCENARIOS AND TACTICS) IS AVAILABLE AND A LIST OF SUCH INFORMATION (INCLUDING SECURITY CLASSIFICATIONS) HAS BEEN PROVIDED TO COMOPTEVFOR.

- a. Does COMOPTEVFOR have the most current threat assessment from the intelligence community?
- b. Was additional or updated threat data obtained from the Scientific Intelligence Liaison Officer (STILO)?
- c. Has there been a change in the threat that could jeopardize OPTEVFOR's assessment of the system?
- d. Has there been a change in the threat that may warrant change of operational/performance thresholds or test scenarios?

13. THE SYSTEM SAFETY PROGRAM HAS BEEN SATISFACTORILY COMPLETED.

- a. Has the system safety program been developed in accordance with MIL-STD-882?
- b. Has the system been engineered to minimize the risk of operator error, and to reduce the impact of error on system and operator safety?
- c. What are the particular hazards to the operator and maintainer of the system (life support, electrical shock, etc.) and have these hazards been reduced to an acceptable level? How are these documented?

- d. Does the system/equipment have approval from the Weapon System Explosive Safety Review Board, if required?
- e. Has a hazard analysis been performed?
- f. Have all hazards been identified?
- g. Have the identified hazards been corrected?
- h. Can the system be safely operated and maintained in its intended environment?
- i. Have the operators been trained to perform emergency procedures?
- j. Are all required emergency resources available to support testing?
- k. Are appropriate emergency rescue measures available to mitigate personnel injury?
- l. Have required certifications for diver life support systems been obtained from SEA 00C?

14. THE SYSTEM COMPLIES WITH NAVY OCCUPATIONAL SAFETY AND HEALTH/HAZARDOUS WASTE REQUIREMENTS WHERE APPLICABLE.

- a. Has the program manager conducted an assessment of the environmental effects of the program either through an EA (Environmental Assessment) or a PEA (Programmatic Environmental Analysis) as outlined in OPNAVINST 5090.1? If so, are there any environmental restrictions to conducting the test at the OT site? Have these restrictions been documented and forwarded to CNO N45?
- b. Will any Environmental, Safety or HAZMAT restrictions prevent OPTEVFOR from resolving Critical Operational Issues (COI's) or conducting the test?
- c. Will the system be used differently in OT that would require new Safety and Hazardous waste procedures or an Environmental Assessment?
- d. Has coordination with the test site/platform in the safe handling, loading, use, disposal and emergency procedures for the system been performed?

15. SOFTWARE MATURITY METRICS ANALYSIS DEMONSTRATE THE SOFTWARE IS STABLE AND EXPECTED TO PERFORM AT A LEVEL COMMENSURATE WITH THE OPERATIONAL TEST PHASE.

- a. How mature is the software?
 - (1) Are the software requirements stable? How many ECP's have been written against requirements over the course of the test program?
 - (2) Is the software design stable? What percentage of the software code/modules have been modified over the course of the test program?
- b. Is the software test program complete?
 - (1) To what degree have the number of software tests been successfully completed compared to the total number of software test requirements?
 - (2) To what degree have the software paths or decision points been tested compared to the total number of paths or decision points?
- c. Will the quality of the software support the operational test?
 - (1) What are the number of software faults/reports generated and resolved during testing by priority, age and status?

16. FOR SOFTWARE QUALIFICATION TESTING (SQT), A STATEMENT OF FUNCTIONALITY, DESCRIBING THE SOFTWARE CAPABILITY, HAS BEEN PROVIDED TO OPTEVFOR.

- a. Does the statement of functionality address all of the latest changes?
- b. Does OPTEVFOR agree that the statement adequately describes the software being presented for test?

17. FOR PROGRAMS EMPLOYING SOFTWARE, THERE ARE NO UNRESOLVED PRIORITY 1 OR 2 PROBLEM REPORTS (SPR), AND ALL PRIORITY 3 PROBLEMS ARE DOCUMENTED WITH APPROPRIATE IMPACT ANALYSIS.

- a. Have all PRIORITY 1 and 2 ("High Level") PTR's been resolved?
 - (1) If not, has OPTEVFOR and CNO been made aware of this?
 - (2) How will any PTR's be corrected before the start of OT?
 - (3) What regression testing will be done to verify fixes?
 - (4) Can the PTR be repeated in the Lab and in the field?
- b. Has OPTEVFOR and CNO concurred with the impact analysis?
 - (1) For those that cannot be fixed before OT&E, what is the probability of occurrence during OT&E?

Appendix B T&E Manager's Tool Box

The Law

- TITLE 10 USC 2399 (Major/oversight Programs)
 - AN INDEPENDENT PHASE OF OPERATIONAL TEST AND EVALUATION SHALL BE CONDUCTED BEFORE PROCEEDING BEYOND LRIP. *(DOES NOT MEAN THAT ALL OT HAS TO BE TOTALLY INDEPENDENT)*
 - THE OPERATIONAL TEST AGENCY SHALL INDEPENDENTLY DETERMINE THE QUANTITY OF ARTICLES FOR TESTING. *(BUT PM HAS TO DETERMINE IF IT IS EXECUTABLE)*
 - OT TEST PLAN MUST BE APPROVED IN WRITING BEFORE PROCEEDING INTO OPERATIONAL TEST.
- TITLE 10 USC 2366 (Covered Systems And Major Munitions)
 - COMPLETED LFT&E AND CORRECTIONS BEFORE PROCEEDING BEYOND LOW-RATE INITIAL PRODUCTION

Everything else in the DOD 5000.2 and SECNAVINST 500.2
is negotiable!

TIME TESTED T&E POLICIES

- ➡ T&E must be **tailored** to the risks and fiscal limitations of each program.
- ➡ T&E must be **planned** to support *both* DT and OT objectives where possible.
- ➡ Integrated T&E working groups critical to success.
- ➡ The TEMP **identifies** system capabilities and limitations, particularly in test parameters and thresholds.
- ➡ T&E must **support** development and engineering design, evaluate logistics supportability, and verify attainment of performance requirements.
- ➡ Systems will proceed to OPEVAL only when ready. TECHEVAL will be conducted as a demonstration of the system's readiness for OPEVAL and Fleet Intro.

Major Players In TEMP and T&E

- **PM**
 - Proposed MOE/MOS Thresholds, well defined, tied to the ORD.
 - Realistic and minimum acceptable Critical Technical Parameters (CTPs) traceable to the MOE/MOSs
 - A test schedule linked to the program schedule
 - Affordable DT test resource requirements
- **OPTEVFOR**
 - MOE Parameters selection
 - A realistic Part IV tied to the program
 - Affordable and available OT test resource requirements
- **CNO (N7/N912)**
 - Support for translating ORD reqs into TEMP criteria
 - Validation of the proposed MOE/MOS thresholds
 - Staffing of your TEMP through ASN/OSD

Major Players(cont.)

- **OPTEVFOR**
 - Independent, focus is on operational impact
 - Executes OT per the TEMP
 - Primary input to TEMP Part IV and OT portion of Part V
- **ASN (RDA)**
 - Minor player in T&E planning
 - Focus is on matching APB to TEMP parameters
- **OUSD (DT&E)**
 - OSD DT Representative on IPT for major/oversight programs
 - Focussed on using DT to answer the risk question
 - Emphasis on using test sites and M&S efficiently
- **OUSD (DOT&E)**
 - OSD OT Representative on IPT for major/oversight programs
 - Use OT to answer global warfighting value added questions
 - Emphasis on test realism and excursions using operational forces

DOD 5000.2 and T&E

- Integrated T&E, Combined DT/OT The Norm
- More Testing Up Front Before Program Initiation (ATD/ACTD/Other Demos)
- No Single System End-State, Incremental Performance Baselines
- Each Baseline Increment Defined and Tested In OT
- Milestone C Before OPEVAL, But “Informal” Review Before Fleet Release
- Completed LFT&E and Corrections Before Rate production and Deployment
- Continued Oversight To Evaluate “Significant” Changes To System Or Use In “Substantially New Environment”

SECNAV 5000.2 and T&E

- * OT Cert Criteria emphasize software maturity
 - * software metric goals met
 - * no High Priority STR's
- * OPEVAL is required for full Fleet release of Software to a new platform
- * Separation between TECHEVAL and OPEVAL
 - * 30 days desired, but usually only 1 week or less
- * Any Environmental Analysis required must be completed before **any** DT or OT
- * Stresses T&E IPTs with all players

The TEMP Today

- ➡ ***Still mandatory*** and the primary ***contract*** between the PM and operational testers.
- ➡ ***Translates*** general ORD requirements and thresholds into manageable and testable parameters.
- ➡ ***Interprets*** system capabilities and limitations, particularly in test parameters and thresholds to reduce ambiguities
- ➡ The primary vehicle by which OSD conducts T&E ***oversight*** for selected programs.
- ➡ One of the few acquisition documents supporting more than the Milestone Decision review.
- ➡ Major TEMPs take about 6-14 months to develop, 1-3 months to approve.

SPECTRUM OF OPTEVFOR SERVICES

- ***Early Operational Assessment (EOA)***
 - Formal risk assessment report to support pre Milestone II/B decisions where no product exists.
- ***DT Assist***
 - Observe DT and provide operator perspective input. No formal report
- ***Operational Assessment (OA)***
 - Assessment of systems during development to support LRIP. Formal report.
- ***Combined DT/OT Event***
 - Reducing the scope of OT by formally combining some DT and OT events. Part of the OPEVAL database.
- ***Operational Evaluation (OPEVAL)***
 - Final formal independent phase of OT for Milestone III

T&E Management Lessons

- T&E manager is the IPT lead. Don't give in to group think.
- TEMP is your executable contract, don't include items you can't live with.
- Beware thresholds not backed up by mission analysis.
- Expect OPTEVFOR to request funds for travel and contractor support.
- Recognize the limits of OPTEVFOR to witness or conduct extended tests.
- Evolutionary development and test are best for software systems
- Highest risk hardware environmental tests are those complying with Hi-Impact Shock and Vibration.
- Highest risk software tests are those during integration with existing ship systems.
- Insist on complete test reports from the TDA
- "Production Representative" does not necessarily mean a production unit.
- Always review the OT test plan.
- Although we test to learn, an OPEVAL is still assessed as "pass-fail".

T&E and Acquisition Reform

- ✓ Facilitates use of existing contractor/fleet/agency data vice new. But....quality and usefulness requires assessment.
- ✓ Combining Industrial, Certification, DT and OT tests shortens schedule. But....requires more management attention and risk.
- ✓ COTS shortens development. But....does not necessarily mean less T&E.
- ✓ Performance Based specs offer unique solutions. But....system tested at OT may not be system fielded.
- ✓ Using M&S to answer questions before building prototypes. But....may get less EMD units for T&E.
- ✓ M&S being used to much greater extent to answer T&E questions. But....requires validation to use.

AR gets equipment faster, but they must still be tested!

T&E Tools

- **Guidance**

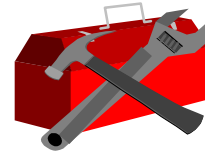
- NAVSEA Systems Acq T&E Management Handbook (2000 Version to be released)
- DSMC T&E Management Guide
- COMOPTEVFOR OTD Guide; www.cotf.navy.mil
- Defense Acquisition Deskbook ; www.deskbook.osd.mil

- **Training**

- DAU TST 101, 201, 301
- OPTEVFOR OTD Course

- **T&E Management Tools**

- Joint Investment in Systems, Technology, Test and Training (JIST3) www.jcte.jcs.mil
- DOD's Automated Test Planning System (ATPS) www.atps.saic.com



**NAVSEA T&E homepage at
www.navsea.navy.mil/navyseate**

OT CERTIFICATION CRITERIA

1. THE TEMP IS CURRENT AND APPROVED.
2. DT THRESHOLDS HAVE BEEN MET, ALL TEST REPORTS PUBLISHED AND DISTRIBUTED.
3. ALL DESIGN PROBLEMS IDENTIFIED AND CORRECTIVE ACTIONS IN PROCESS.
4. SYSTEM OPERATING AND MAINTENANCE DOCUMENTS DISTRIBUTED TO OPTEVFOR.
5. ADEQUATE LOGISTIC SUPPORT IS AVAILABLE AS DOCUMENTED IN THE TEMP
6. SYSTEM TECHNICAL DOCUMENTATION PROVIDED TO OPTEVFOR.
7. THE OT&E MANNING IS THE SAME PLANNED FOR FLEET UNITS.
8. THE NAVY TRAINING PLAN HAS BEEN APPROVED.
9. TRAINING COMPLETED AND IS REPRESENTATIVE OF THAT PLANNED FOR FLEET UNITS.
10. ALL RESOURCES REQUIRED FOR OPERATIONAL TEST ARE AVAILABLE.
11. THE SYSTEM IS PRODUCTION REPRESENTATIVE.
12. ALL THREAT INFORMATION PROVIDED TO OPTEVFOR.
13. THE SYSTEM SAFETY PROGRAM HAS BEEN SATISFACTORILY COMPLETED.
14. SYSTEM COMPLIES WITH NAVY SAFETY AND HEALTH/HAZARDOUS WASTE REQs.
15. SOFTWARE MATURITY METRICS SHOW THE SOFTWARE IS STABLE.
16. STATEMENT OF FUNCTIONALITY PROVIDED TO OPTEVFOR FOR SOFTWARE QUALIFICATION TESTING (SQT).
17. THERE ARE NO UNRESOLVED SOFTWARE PRIORITY 1 OR 2 PROBLEM REPORTS.

Appendix C

Index

3

3-M · 5-9, A-3, 2

A

AAP · 1-1, 3-5
Abbreviated Acquisition Programs · See AAP
ACAT · 1-1, 2, 5, 6, 7, 9, 2-3, 5, 6, 17, 3-1, 2, 5, 6, 4-1, 2, 3, 4, 6, 7, 5-1, 6, 6-2
ACAT III · 2-4, 6, 3-1, 3
ACAT IV · 2-4, 3-1
accreditation · 2-7, 8, 5-5
Acquisition Categories · See ACAT.
Acquisition Program Baseline · See APB.
ACTD · 1-2, 2-1, 2, 4-1, 2
Advanced Concept Technology Demonstrations · See ACTD.
Advanced Technology Demonstrations · See ATD.
Analysis of Alternatives · See AOA.
Ao · 4, 12, 14, 8, 2
AOA · 2-2, 4-1, 3, 7
APB · 2-1, 4, 3-2, 7, 4-4, 5
ARCI · 1-8
ASN (RDA) · 1-1, 4, 2-3, 3-1
Assistant Secretary of the Navy for Research Development and Acquisition · See ASN(RDA).
ATD · 1-2, 4-1
Availability · 2-13

B

burn-in test · 6-2

C

C4I · 1-13, 2-8, 9, 4-4
C4ISR 1-13, 14
CAD · 4-1, 2
CAIV · 2-4
CE · 4-1, 2, 3, 4
Certification Criteria · 5-8
Certification of Readiness · 5-8, 11, A1
Chief of Naval Operations · See CNO.
CINCLANTFLT · 1-6
CINCPACFLT · 1-6
CNO · 1-1, 6, 9, 10, 15, 16, 17, 2-1, 2, 3, 5, 7, 3-3, 5, 6, 7, 8, 10, 11, 12, 4-1, 5, 6, 7
CNO N912 · 1-1, 17, 3-5, 8, 10, 11

Coast Guard · 2-6
COATS · 2-10, 11
COI · 3-4, 5, 5-9, A6
Combat Control Systems · 1-12, 13
Combat System Engineering Development Site · See CSEDS.
combined DT/OT · 2-5, 8
Command and Control System Module Off-Hull Assembly and Test Site · See COATS
Command Control Communications Computers and Intelligence · See C4I
Commander, Operational Test and Evaluation Force · See OTF/COTF.
Commercial-Off-The-Shelf · See COTS
COMOPTEVFOR · 1-4, 5, 6, 2-1, 2, 5, 3-4, 6, 7, 4-4, 5, 8, 9, 10, 11, 5-1, 2, 3, 4, 6
Component Advanced Development · See CAD.
Computer Software Testing · 2-14
Concept and Technology Development · 1-2, 4-2, 4-5
Concept Exploration · See CE
Contractor Test and Integration Facilities · 2-11
Cost as an Independent Variable · See CAIV
COTS · 1-1, 6, 7, 9, 14
covered system · 1-10
Critical Operational Issues · See COI.
Critical Technical Parameters · 2-3, 5, 14, 3-3
CSEDS · 2-10, 5-6
CTP · See Critical Technical Parameters.

D

DA · 1-4, 5-10, 11
De-certification · 5-13
Defense Acquisition Desk Book · See Deskbook
DEP · 2-10, 11, 12
Deskbook · 2-14
developing agencies · See DA
Developmental T&E · See DT and DT&E
Distributed Engineering Plant · See DEP.
DOD · 1-2, 4, 5, 9, 10, 15, 2-2, 6
DoD 5000.2 · 1-18, 2-2
DOT&E · 2-10, 12, 3-1, 4, 5, 10
DT · 1-4, 5, 6, 7, 8, 9, 2-1, 2, 3, 5, 6, 7, 8, 10, 11, 13, 14, 17, 18, 3-1, 3, 4, 5, 6, 7, 4-3, 4, 6, 7, 8, 9, 5-1, 2, 4, 5, 6, 8, 9, 10, 11, 12, 13, A-1, 2
DT&E · 1-4, 5, 9, 2-1, 2, 3, 5, 10, 3-1, 3, 4, 3, 7, 4-4, 5, 5-1, 5, 6, 8, 10, A-1, 2
DT/OT · 2-5, 10, 5-17, A-1

E

EA · 2-16, A-6
Early Operational Assessments · See EOA.
EDM · 1-2, 3, 5-1, 2, 5

Emergent Requirements · 2-17
EMI · 2-14, 6-3, A-2
Engineering Development Models · See EDM.
Environmental Assessments · See EA
Environmental qualification Testing (EQT)
2-13
Environmental Regulations · 2-9
EOA · 1-6, 4-2, 4, 5
Evolutionary acquisition · 1-8
Explosives 1-10

F

Fault Profiles · 2-15
Finding Of No Significant Impact
FONSI · 2-16
Fleet Battleforce Experiments · 1-2
Fleet Services Requests 2-16
Follow-on Operational Test and Evaluation · See
FOT&E.
FONSI · 2-16
Foreign Weapons Evaluation · See FWE
FOT&E · 1-4, 5, 9, 5-2, 5, 6, 9, 10, 11, A-1, 3, 4,
5, 6
FRP · 2-17, 3-3, 6, 8, 5-1, 2, 3, 5, 10, 13, 15, A-1
Full Rate Production · See FRP.
funding · 1-2, 6, 9, 2-6, 16, 3-2, 3, 6, 7, 9, 5-12, A-
2
FWE · 1-2, 2-17

G

Government Test and Integration Facilities · 2-11

H

Harmonizing DT and OT · 2-6
HERO Tests · 1-10
HM&E · 1-14, 15, 4-6

I

IEEE 12207 · 2-15
ILS · 1-1, 7, 10, 5-10
Initial Operational T&E · See IOT&E
installation · 1-9, 2-2, 6, 9, 10, 11, 17, 3-4, 6, 5-1,
5, 6, 12, A-1, 2, 3, 4
Integrated Logistics Support · See ILS.
Integrated Process Teams · See IPT. See IPT
Integrated Test and Evaluation Plan · See ITEP
Integrated Schedule · 3-4
Interim Progress Review · See IPR.
Interoperability · 2-7, 8, 11, 3-2, 5, 4-2, 3, 5, 8, 5-
2, 3
IOT&E · 1-4, 5, 10, 11, 6, 3-3, 5-1, 5, A-1
IPR · 3-6, 4-3, 5-2, 3

IPT · 2-12, 3-1, 5, 6, 4-5, 5-9
ITEP · 3-5, 5-2

J

JIST3 · 2-9

K

Key Performance Parameters · See KPP.
KPP · 2-4, 4-7

L

land-based test site · See LBTS
LBTS · 2-9, 10, 11, 5-1, A-4
Lethality · 1-10, 16
LFT&E · 1-2, 4, 6, 9, 10, 16, 3-1
Live-Fire T&E · See LFT&E.
Low Rate Initial Production · See LRIP.
LRIP · 1-2, 2-10, 11, 18, 4-3, 8, 9, 5-1, 5, 6

M

M&S · 2-6, 7, 8, 4-1, 4, 5, 7, 5-6, 8
Maintainability · 2-3, 4
maintenance assist modules · See MAM.
Maintenance Data System · See MDS
MAM · 2-12
MDA · 1-2, 6, 2-18, 3-1, 5-2
MDS · 6-3
mean-time-to-repair · See MTTR.
Measures of Effectiveness · See MOE.
Measures of Performance · See MOP.
Measures of Suitability · See MOS.
Milestone A · 1-2, 3-1, 4-1, 2
Milestone B · 1-2, 3, 3-1, 2, 3, 4, 5, 6, 7, 8, 4-1, 5
5-1
Milestone C · 4-2, 3, 5-1, 2, 6
Milestone Decision Authority · See MDA.
Milestone Reviews · 1-2, 4-4, 7
MIL-SPEC · 1-1
MIL-STD · 1-10 2-13, 14, A-6
modeling and simulation · See M&S.
MOE · 2-3, 4, 3-2, 4, 5, 7
MOP · 2-3, 4, 5
MOS · 2-3, 5, 2-14, 3-2, 4, 7, 4-6
MOT&E · 4-3, 5-6
MTTR · 2-12, 4-6, 8

N

NAPDD · 2-3
National Environmental Policy Act · See NEPA

NAVSEA · 1-1, 4, 6, 2-1, 5, 12, 15, 17, 3-1, 5, 7,
4-2, 3, 8, 4-7, 9, 12, 13, 5-3
NDI · 1-1, 6, 9, 4-1, 3, 4, 7
NEPA · 2-15
NIB · 2-17
Non-Acquisition Program Description Document ·
See NAPDD
Non-Developmental Item · See NDI.
not-to-interfere basis · See NIB.

O

OA · 1-6, 7, 3-3, 5-1, 3
Observation of Operational Capability · See OOC
Office of the Secretary of Defense · See OSD
OOC · 1-6
Operational Assessment · See OA.
Operational Availability · See Ao.
operational effectiveness · 1-4, 5, 6, 8, 9, 2-1, 4-4,
5-1, 6, 11
Operational Requirements Document · See ORD.
operational suitability · 1-4, 5, 8, 2-1, 15, 3-7, 5-1,
6, 11
Operational Suitability Guide · 2-14
Operational Test and Evaluation · See OT&E.
Operational Test Director · See OTD.
Operational Test Readiness Review · See OTRR
Operational Testing · See OT
Operations Security · See OPSEC
OPEVAL · 1-1, 4, 5, 6, 8, 9, 2-3, 5, 6, 10, 11, 13,
14, 15, 3-3, 4, 6, 7, 4-3, 8, 5-1, 2, 4, 5, 6, 7, 8,
9, 10, 11, 12, 14, 15, A-1, 3, 4, 5, 6
OPNAV · 1-1, 4, 2-2, 3, 4, 5, 17, 3-2, 5, 6, 7, 4-5,
6, 7, 8, 10, 12, A-1, 4
OPNAVINST 5090.1 · 2-15, A-6
OPSEC · 2-5, 5-11, 13, A-5
OPTEVFOR · 1-5, 6, 7, 9, 2-2, 3, 5, 8, 10, 11, 14,
16, 17, 18, 3-1, 3, 4, 5, 7, 4-1, 2, 3, 4, 5, 6, 7,
8, 5-1, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, A-1, 2,
4, 5, 6, 7
ORD · 2-1, 3, 4, 17, 3-1, 2, 6, 7, 4-1, 2, 3, 5, 6, 7,
8, 9, 11
OSD · 1-1, 6, 9, 10, 12, 3-1, 4, 7, 5, 10
OT · 1-1, 4, 5, 6, 7, 8, 9, 2-1, 2, 3, 4, 5, 6, 7, 10,
11, 13, 14, 16, 17, 18, 3-1, 3, 4, 5, 6, 4-3, 4, 6,
7, 8, 5-1, 4, 6, 7, 8, 9, 10, 11, 12, 13, A-1, 2, 3,
4, 5, 6, 7
OT&E · 1-1, 4, 5, 9, 2-1, 2, 5, 10, 16, 3-1, 3, 4, 5,
4-3, 4, 7, 5-1, 2, 4, 5, 8, 9, 10, 11, A-1, 2, 3, 4,
5, 7
OTD · 1-6, 2-2, 3, 7, 3-3, 4, 4-2, 6, 7, 10, 11, 12,
5-8, 14
OTRR · 5-11, 12, 13
Oversight · 1-7, 4-1

P

PAT&E · 6-2, 3
PEO · 1-4, 6, 3-1, 2, 5-5, 6, 8, 9, 10, 11, 6-1

Piggy-Back testing · 2-5
Pk · 1-10, 11
Planned Maintenance System · See PMS.
Planning, Programming and Budgeting System ·
See PPBS
PM · 1-1, 6, 7, 9, 2-10, 4, 5, 8, 12, 2-13, 14, 15,
16, 3-2, 4, 5, 6, 4-4, 1, 5-1, 6, 7, 11
PMS · 5-10, 6-3, A-2
PPBS · 4-7, 5-12
PPP · 2-16
Priority ONE · 2-17
Probability of Kill · See Pk
Production Acceptance Test and Evaluation ·
See PAT&E.
Program Executive Officers · See PEO
Program Protection Plan · See PPP
Progress Reviews · 3-4
Proof testing · 6-3
PTR · 5-3

Q

quick look report · 5-5, 13

R

R&D · 1-1, 2, 4, 5, 6, 8, 9, 2-1, 2, 17, 7, 6-1
re-certification · 5-12
Reliability Maintainability Availability (RMA) · 2-13
Reports · 3-1, 5-9
Research and Development · See R&D

S

sampling test · 6-2
SAR · 3-6
screening tests · 6-2
SDD · 2-3, 4, 4-1, 2, 4
SECNAVINST 5000.2 · 2-9, 15, 18, 3-2, 5-8, 14
SECNAVINST 5000.2B · 1-2, 4-5, 7, 5-8, 10, 11,
6-2
Selected Acquisition Report · See SAR
Sensor Systems 1-12
Ship Acquisition Program · 1-15
Ship Construction Tests and Trials Manual · 6-3
SIM-STIM · 1-11
Software Development · 2-15
software qualification testing · See SQT
Software Upgrade Program · 1-14
SQT · 5-11, A-7
Summary Report · 3-6
surrogate · 1-9, 10, 2-12, 3-4, 7
System Development · 1-8, 2-1, 4-1, 2, 3, 8, 9, 5-
1, 2, 3, 5
System Development and Demonstration · See
System Development.
System Integration · 5-2
systems engineering · 1-4, 5, 2-10, 4-3

T

T&E · 1-1, 2, 4, 5, 6, 7, 8, 9, 10, 2-1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 15, 16, 17, 18, 3-1, 2, 4, 5, 6, 7, 4-1, 2, 3, 4, 5, 6, 7, 8, 9, 5-1, 2, 4, 5, 6, 7, 9, 10, 11, 12, 13, 6-1, 3, A-1
TDA · 2-2, 4, 4-1, 2
TECHEVAL · 1-5, 6, 8, 9, 2-4, 5, 6, 10, 11, 13, 14, 3-4, 6, 7, 4-3, 5-1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, A-1, 2, 3, 5
technical design agent · See TDA.
Technology · 1-1
TEIN · 2-17, 3-5, 6, 4-3
TEMP · 1-4, 6, 9, 2-1, 2, 3, 4, 5, 10, 14, 17, 18, 3-1, 2, 3, 4, 5, 6, 7, 8, 4-2, 3, 4, 5, 6, 7, 8, 5-2, 8, 9, 11, 12, 13, 6-1, 2, A-1, 2, 3
Test and Evaluation Identification Number · See TEIN.
Test and Evaluation Master Plan · See TEMP.
test federations · 2-6
Test Planning Working Group · See TPWG. See TPWG
Test Readiness Review (TRR) · 5-7
Test Site Selection · 2-9
Testability · 2-12, 13
Threat Assessments · 1-4, 3-2
Threat Target Validation (TTV) · 2-12
threshold · 1-4, 5, 14, 15, 2-4, 5, 3-7, 4-3, 5, 6, 7, 8, 6, 5-2
Title 10 US Code Section 2366 · 1-1

Title 10 US Code Section 2400 · 1-1, 9
TPWG · 2-6, 5-6
Traceability · 2-15
traceability matrix · 2-3

U

Under Secretary of Defense · See OSD
United States Code · See USC
USC 2366· 1-15
USC 2399· 1-1, 2-17, 5-6

V

Validation, Verification and Accreditation · See VV&A
Vulnerability· 1-10, 16
VV&A · 2-8, 3-5

W

warshot · 1-11
Weapon System Explosive Safety Review Board (WSESRB)· 1-10, 5-13